

crff 163903

Jarrell, Noble

From: Ramirez, Delia
Sent: Thursday, August 25, 2005 3:12 PM
To: Jarrell, Noble
Subject: FW: 09/598,982

-----Original Message-----

From: Ramirez, Delia
Sent: Monday, August 01, 2005 6:14 PM
To: Schreiber, David
Subject: 09/598,982

Hi,

I would like to request the following interference searches:

1. SEQ ID NO:52 in the protein and nucleic acid databases (commercial and interference)
2. SEQ ID NO:20 and 21 in the nucleic acid databases (interference)
3. SEQ ID NO:21 in the protein databases (interference)
4. an alignment of SEQ ID 9, 11, 21, 23, 25, 27, 37, 39, 41 and 43
5. an alignment of SEQ ID NO: 8, 10, 20, 22, 24, 26, 36, 38, 40, 42

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

Noble 8/31/05
Compugn 10/02
3NA 2MA 10/04
Commercial
interfer.

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 21:12:24 ; Search time 607 Seconds
(without alignments) 2684.068 Million cell updates/sec

Title: US-09-598-982C-21
Perfect score: 1393
Sequence: 1 LERKIVGGQEARPRKRPWQV.....IYTRVYTYLDMHHVYPRKP 249

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dgapop 6.0 , Dextext 7.0

Searched: 7311713 seqs, 3271544945 residues
Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n_model -DEV=xlp
-Q=/cgn2_1/USPRO_spool_P/US0959822/runat_25082005_165520_304/app_query_faetra_1.391
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnmp -MINMATCH=0.1
-LOOPCL=0 -IOPEXT=0 -UNITS=bits -SPARR=1 -END=-1 -MATRIX=blowm62
-TRANS=human40 cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0959822@cgn2_1_723@runat_25082005_165520_304
-NCPU=6 -ICPU=3 -NO MMAP -LARGEOBJECT -NEG SCORES=0 -WAIT -DSPBLIOCL=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA :
1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubna/FCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubna/US10D_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubna/US10E_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubna/US10G_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubna/US10H_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubna/US10I_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubna/US10J_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB, ID, Description. Contains 45 rows of search results.

ALIGNMENTS

RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Carroll, Joseph M.
APPLICANT: Healy, Aileen S.
APPLICANT: Welch, Louise M.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
15402, 340, 10217, 857, 1761, 8990 OR 13249 MOLECULES
FILE REFERENCE: MPI02-019P1R0NMNM

Pred. No. is the number of results predicted by chance to have a

```

CURRENT APPLICATION NUMBER: US/10/352,684A
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 60/354,333
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: US 60/360,258
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/364,476
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/375,626
PRIOR FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: US 60/386,494
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/390,965
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: US 60/392,480
PRIOR FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US 60/394,128
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/399,783
PRIOR FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: US 60/403,221
PRIOR FILING DATE: 2002-08-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 1143
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (6)...(833)
US-10-352-684A-47

Alignment Scores:
Pred. No.: 2,05e-158 Length: 1143
Score: 1368.00 Matches: 244
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 98.21% Indels: 0
DB: 17 Gaps: 0

US-09-598-982c-21 (1-249) x US-10-352-684A-47 (1-1143)
QY 5 IlevaIGlyGInGInUaIaProArGserLySTrPrOTrPInVaISerLeuArGVal 24
Db ATCGTTGGGGGTCCAGAGAGCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAGTC 155
QY 25 HnSGlyPrOTyTTrPMethiSPhEcYsgIyGlySerLeuIleHiSProGInTrpValLeu 44
Db CAGGGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 215
QY 45 ThraIaAlaIaCyAvVaIaGlyProAspVaIlyVaAspLeuIaIaIaIaIaIaIaIaIaIa 64
Db ACCGGAGCCCACTGCTGGAGCCGGACCGTCAAGAGATCTGGCCGCTTCAAGGTTGCACTG 275
QY 65 ArGdIuGInHnIbVenuTyTrYrGInIaSPrgInLeuLeuProVaISerArGIIeIleVaIhIS 84
Db CCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 335
QY 85 ProGInPhETyTThraIaGInIleGlyIaAaSPrlaIaIaIaIaIaIaIaIaIaIaIaIaIa 104
Db CCAAGTATTACACCCCGAGATCGGAGGAGCATGCGCCCTGTGAGCTGGAGGAGCGCG 395
QY 105 VallyVaISerSerHisVaIhISThraValIThrLeuProAlaSerGIuThrPhEPro 124
Db GTGAAGTCTTCCAGCCAGCTCCACAGCGTCAACCCCTGCTTCAAGAGCCTTCCCTCC 455
QY 125 ProGlyMeCProCySTrPValIhrgIyTrPglYAsPValaSPaSnAspGIuArGLeuPro 144
Db CCGGGATCCCGTGGTGGGATCACTGGCTGGGGGAGATGTGACAAATGATAGAGCCCTCCA 515
QY 145 PropProPheroleuLySgInVallyIySValProIleMetGIuAsnHIStleCyAsaPaIa 164

```

```

Db CCGCCATTTCTCTCAACAGGTGAAGGTCCCCCAVAAAGAAACCAACATTTGTACCGCA 575
QY 155 LysTYrHISleUGlyAlaATyTThrGlyVaSPaPValaRGIleValaRAspAPMeTLeu 184
Db AAATACCACTTTGGCCCTTACCGGAGACAGACGTCCCAATGCTCCGTGACAGCATGCTG 635
QY 185 CysAlaGIaSnThraRArgAAspSerCySgInGIyaSPserGIyGIyProleuValCys 204
Db TGTGCCGGGAACACCCCGAGGAGACTCATGCCAGAGGCGACTCCGGAGGGCCCTGGTGTGC 695
QY 205 LysVaIaSnGIyThrTrPLeuGInIaGlyValaISerTrPGLyGInGlyCySAIaGIn 224
Db AAGGTGAATGGCACCTGGCTGAGGCGGGCGGTCAAGCTGGGGGAGGGGCTGTGCCAG 755
QY 225 ProAenArPProGlyIleTyTThraRgValIThrTYrTrLeuAspTrpIleHiSHISrYr 244
Db CCNAACCGGCTGGAGATGACCCCGTGTCACTTACTTGGAGCTGGATCCACCACTAT 815
QY 245 ValProlYsLySPro 249
Db GTCCCAAAAAGCCG 830

RESULT 2
US-10-287-226-93
Sequence 93, Application US/10287226
Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Beighs, Constance,
APPLICANT: Boldog, Ference,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: Dipippo, Vincent A.,
APPLICANT: Edinger, Shlomoit R.,
APPLICANT: Eilsen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Eshe A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khrantsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malyankar, Uziel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Murali dhara,
APPLICANT: Patburajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vermet, Corine A.M.,
APPLICANT: Zernusen, Bryan D.,
APPLICANT: Zhong, Mei

TITLE OP INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287, 226
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04

```

PRIOR APPLICATION NUMBER: 60/360,148
 PRIOR FILING DATE: 2002-02-27
 PRIOR APPLICATION NUMBER: 60/364,000
 PRIOR FILING DATE: 2002-03-13
 PRIOR APPLICATION NUMBER: 60/404,821
 PRIOR FILING DATE: 2002-08-20
 PRIOR APPLICATION NUMBER: 60/334,526
 PRIOR FILING DATE: 2001-11-30
 PRIOR APPLICATION NUMBER: 60/354,409
 PRIOR FILING DATE: 2002-02-04
 PRIOR APPLICATION NUMBER: 60/364,227
 PRIOR FILING DATE: 2002-03-13
 PRIOR APPLICATION NUMBER: 60/334,027
 PRIOR FILING DATE: 2001-11-28
 PRIOR APPLICATION NUMBER: 60/331,641
 PRIOR FILING DATE: 2001-11-20
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 673
 SOFTWARE: Cursesqlist version 0.1
 SEQ ID NO: 93
 LENGTH: 1145
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (8)..(832)
 US-10-287-226-93

Alignment Scores:
 Pred. No.: 2,06e-158 Length: 1145
 Score: 1368.00 Matches: 244
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 98.21% Indels: 0
 DB: 18 Gaps: 0

US-09-598-982c-21 (1-249) x US-10-287-226-93 (1-1145)

QY 5 ILEVALGIYGLINGLUALAPROARSGERYSTPPTPTGIVALSERIEUARYVA 24
 |||
 DB 98 ATCGTGGGGGTCCAGGAGCCGCCCGAGCAAGTGGCCCTGCGAGGTGAGAGTCC 157
 |||
 QY 25 HISGLYPTQYTRTTRMETHISPHCYSGIYGLYSETLEUIIHHISPROGLNTRPVALLIU 44
 |||
 DB 158 CAGGGCCCATCTGGATGACTTCTGGGGGGCTCCCTCATCCACCAGTGGGTCTG 217
 |||
 QY 45 THRALAAIAALCYEVALGIYPROARPVALLYSAPLEUALAIALEUARYVALGINLEU 64
 |||
 DB 218 ACCGAGGCGCACTGGCTGGGAGCCGACGTCMAAGGATCTGGCCGCTCAAGGTGCAACTG 277
 |||
 QY 65 ARGGLVGINHISLEUYRYTYRGINAPRGINLEUPEVALSERFARGLIEIIEVALHIS 84
 |||
 DB 278 CCGGAGGACGACCTCTACTACCAAGACCACTGCTGCCGATCCAGGATCATCTGTCAC 337
 |||
 QY 85 PROGINPHEVYTRTHRALAGINILEGIYALAAAPLIEALALEUENGLIUNGLIUNP 104
 |||
 DB 338 CCACAGTTTAAACCGCCAGANTCGGAGCGCAATGCGCTGTGGAGCTGGAGGAGGCTG 397
 |||
 QY 105 VALLYVALSERSETHISVAIHISVTHRVALLIHRLEUPROPROALASERGIVLTHRPHEP 124
 |||
 DB 398 GTGAAGGTTCCAGGCAAGTCCACACCGTCAACCCCTGCCCTCAGAGACCTTCCCC 457
 |||
 QY 125 PROGLYMETPCQYSTRPVALLIHRGLYTRPGLIYAPVALLASPHASAPGLIUAIGLEUP 144
 |||
 DB 458 CCGGGGATCCCGTGGCTGCTCACTGGCTGGGGGAGATGAGACAAATGATGAGGCCCTCCA 517
 |||
 QY 145 PROPROPHLEUPEUENGLIUNGLIUNGLIUNGLIUNGLIUNGLIUNGLIUNGLIUN 164
 |||
 DB 518 CCGCCATTTCTCTGAAAGGAGGTGAGGTCCTCCCATTAATGAAACCAACATTTGTAGCCGA 577
 |||
 QY 165 LYSTRYTHISLEUENGLIYALATYTRHTRGIYASAPVALLIARGLIEVALARGHAPSPMETLE 184
 |||
 DB 578 AATATACCACTTGGGGCTTACACGGGAGACGAGCGTCCGATCGTCCGTGACGACATGGCTG 637
 |||

CYSALAGIYAENHTRRARGARASPERCYESLNGIYASPERGIYGLYPROLEUVALCY 204
 |||
 DB 638 TGTGCGGAAACACCCGGAGGACTCATTCAGGACCACTCCGAGGGCCCTGCTGTCG 697
 |||
 QY 205 LYSEVALSNGIYTHTRPLEUENGLIYVALSERTRPGLIYGLIYCYEVALAGL 224
 |||
 DB 698 AAGGTGAATGCACTGGCTGTCAGGCGGGGTGTCAGCTGGGGCGAGGGCTGTCAG 757
 |||
 QY 225 PROANHRPPTQYIYIETRYTRHARGVALIHRVYTRVLEUAPTRPIIHHISITR 244
 |||
 DB 758 CCCAACCGGCTGGCATCTACACCCGCTGCTCACTACTACTGATCGATCCACCACTAT 817
 |||
 QY 245 VALPROLYSVP 249
 |||
 DB 818 GTCCCAAAAAGCCG 832
 |||

RESULT 3
 US-10-956-157-2444
 Sequence 2444, Application US/10956157
 Publication No. US20050118625A1
 GENERAL INFORMATION:
 APPLICANT: Wyeth
 TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
 FILE REFERENCE: 031896-043000 (AM 101081)
 CURRENT FILING DATE: 2004-10-04
 NUMBER OF SEQ ID NOS: 319805
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 2444
 LENGTH: 1148
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-956-157-2444

Alignment Scores:
 Pred. No.: 2,06e-158 Length: 1148
 Score: 1368.00 Matches: 244
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 98.21% Indels: 0
 DB: 21 Gaps: 0

US-09-598-982c-21 (1-249) x US-10-956-157-2444 (1-1148)

QY 5 ILEVALGIYGLINGLUALAPROARSGERYSTPPTPTGIVALSERIEUARYVA 24
 |||
 DB 98 ATCGTGGGGGTCCAGGAGCCGCCCGAGCAAGTGGCCCTGCGAGGTGAGAGTCC 157
 |||
 QY 25 HISGLYPTQYTRTTRMETHISPHCYSGIYGLYSETLEUIIHHISPROGLNTRPVALLIU 44
 |||
 DB 158 CAGGGCCCATCTGGATGACTTCTGGGGGGCTCCCTCATCCACCAGTGGGTCTG 217
 |||
 QY 45 THRALAAIAALCYEVALGIYPROARPVALLYSAPLEUALAIALEUARYVALGINLEU 64
 |||
 DB 218 ACCGAGGCGCACTGGCTGGGAGCCGACGTCMAAGGATCTGGCCGCTCAAGGTGCAACTG 277
 |||
 QY 65 ARGGLVGINHISLEUYRYTYRGINAPRGINLEUPEVALSERFARGLIEIIEVALHIS 84
 |||
 DB 278 CCGGAGGACGACCTCTACTACCAAGACCACTGCTGCCGATCCAGGATCATCTGTCAC 337
 |||
 QY 85 PROGINPHEVYTRTHRALAGINILEGIYALAAAPLIEALALEUENGLIUNGLIUNP 104
 |||
 DB 338 CCACAGTTTAAACCGCCAGANTCGGAGCGCAATGCGCTGTGGAGCTGGAGGAGGCTG 397
 |||
 QY 105 VALLYVALSERSETHISVAIHISVTHRVALLIHRLEUPROPROALASERGIVLTHRPHEP 124
 |||
 DB 398 GTGAAGGTTCCAGGCAAGTCCACACCGTCAACCCCTGCCCTCAGAGACCTTCCCC 457
 |||
 QY 125 PROGLYMETPCQYSTRPVALLIHRGLYTRPGLIYAPVALLASPHASAPGLIUAIGLEUP 144
 |||

Db 458 CCGGGGATGCGCTGCTGGCTCACTGGCTGGGGGCGGATGTGGACATGATGAGCGCCCTCCCA 517
 Qy 145 ProProPheProLeuLysGlnValLysValProIleMerGlnuShiIleCysAspAla 164
 Db 518 CCCCCATTTCTCTGAAAGGAGGTTGAGGTTCCCATATGAAACCAATTTGTGCGCA 577
 Qy 165 LysTyrHisIleuGlnValTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
 Db 578 AAATACCACTTGGCGCTGACCGGGAGACGATCGTCCGATCGTCCGATCGCAATGCTCG 637
 Qy 185 CVALAGLYASMTHTARGAAGASPSERCYSGINGLYASPSERGLYGLYPROLEUVALCYS 204
 Db 638 TGTGCGGGGAACCCGGAGGAGCATATGCGAGGACCTCCGGAGGGCCCTGGGTGTC 697
 Qy 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
 Db 698 AAGGTGATGACCTGGCTGGCGAGGGGGGGGTGTGACCTGGGGCGAGGGCTGTGCCCG 757
 Qy 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrIleuAspTrpIleHisIleTyr 244
 Db 758 CCCAACCGGCTGCAATCTACCCCGGTGCTACTACTGTGGACTGGATCCACCACTAT 817
 Qy 245 ValProLysValPro 249
 Db 818 GTCCCAAAAAAGCCG 832

RESULT 4
 US-10-287-226-91
 ; Sequence 91, Application US/10287226
 ; Publication No. US2004086875A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agee, Michele L.,
 ; APPLICANT: Aisobrook, John P.,
 ; APPLICANT: Berghs, Constance,
 ; APPLICANT: Boldog, Ferenc,
 ; APPLICANT: Burgess, Catherine E.,
 ; APPLICANT: Chant, John S.,
 ; APPLICANT: Chaudhuri, Amitabha,
 ; APPLICANT: DiPippo, Vincent A.,
 ; APPLICANT: Edinger, Shlomit R.,
 ; APPLICANT: Eisen, Andrew,
 ; APPLICANT: Eilemann, Karen,
 ; APPLICANT: Ganggoli, Ssha A.,
 ; APPLICANT: Gorman, Linda,
 ; APPLICANT: Gerlach, Valerie,
 ; APPLICANT: Ji, Weizhen,
 ; APPLICANT: Karamtsou, Nikolai,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Malyanakar, Uriel M.,
 ; APPLICANT: MacDougall, John R.,
 ; APPLICANT: Mezes, Peter S.,
 ; APPLICANT: Miller, Charles E.,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Ooi, Chean Eng,
 ; APPLICANT: Oot, Tatsuna,
 ; APPLICANT: Padigarau, Muralidhara,
 ; APPLICANT: Paturajan, Meera,
 ; APPLICANT: Rastelli, Luca,
 ; APPLICANT: Rieger, Daniel K.,
 ; APPLICANT: Rothenberg, Mark E.,
 ; APPLICANT: Shenoy, Suresh G.,
 ; APPLICANT: Spaderna, Steven K.,
 ; APPLICANT: Spytsek, Kimberley A.,
 ; APPLICANT: Taupier, Jr., Raymond J.,
 ; APPLICANT: Vermet, Corine A.M.,
 ; APPLICANT: Zernusen, Bryan D.,

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 FILE REFERENCE: 21402-480C
 CURRENT APPLICATION NUMBER: US/10/287,226
 CURRENT FILING DATE: 2002-11-04
 PRIOR APPLICATION NUMBER: 60/334,421

Db 458 PRIOR FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: 60/354,392
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/360,148
 ; PRIOR FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: 60/364,000
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 60/404,821
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: 60/334,526
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: 60/354,409
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/364,227
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 60/334,027
 ; PRIOR FILING DATE: 2001-11-28
 ; PRIOR APPLICATION NUMBER: 60/331,641
 ; PRIOR FILING DATE: 2001-11-20
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 673
 ; SOFTWARE: Curoseqdist version 0.1
 ; SEQ ID NO 91
 ; LENGTH: 828
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(825)
 ; US-10-287-226-91

Alignment Scores:
 Prd. No.: 5,54e-158 Length: 828
 Score: 1363.00 Matches: 243
 Percent Similarity: 99.18% Conservative: 0
 Best Local Similarity: 99.18% Mismatches: 2
 Query Match: 97.85% Indels: 0
 DB: Gaps: 0

US-09-598-982c-21 (1-249) x US-10-287-226-91 (1-828)

Qy 5 lIeValGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
 Db 91 ATCGTCGGGGGAGCAGAGAGCCCGCCAGGACAAAGTGGCCCTGGCAGGTGAGCCTGAGATC 150
 Qy 25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPArgGlnTrpValLeu 44
 Db 151 CACGGCCCACTTGTGATGCACTTCTGCGGGGCTCCCTCATCCACCCCAAGTGGTCTG 210
 Qy 45 ThrAlaAlaAlaCysValGlyProAspValIlyAspLeuAlaAlaLeuArgValGlnLeu 64
 Db 211 ACCGACGGCACTGCGTGGGACCGGACGTCAAAGATCTGGCCCGCCCTCAGGGTCACTG 270
 Qy 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
 Db 271 CCGGAGCGCACCTTACTTACCGAGCAGCGTCTGCCGGTCAAGAGATCATCTGTGCAC 330
 Qy 85 ProGlnPheTyrThrAlaGlnIleGlyValAlaAspIleAlaLeuLeuGlnLeuGlnIleuPro 104
 Db 331 CCAAGTCTTACACCGCCCAAGTCCGAGCGGACATCCGCTGTGAGCTGGAGAGCCG 390
 Qy 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnThrPhePro 124
 Db 391 GTGAAGCTCTCCAGGCAAGTCCAGCGGTCACCCCTGCCCGCTCAGAGACCTTCCCC 450
 Qy 125 ProGlyMetProCysTyrValThrGlyTyrGlyAspValAspAspArgGlnArgLeuPro 144
 Db 451 CCGGGGATGCGCTGCTGGGTCACTGGCTGGGGCGATGTGCAAAATGATGAGGGCCCTCCA 510
 Qy 145 ProProPheProLeuLysGlnValLysValProIleMetGlnuShiIleCysAspAla 164
 Db 511 CCGCATTTCTCTGAAAGGAGGTTGAGGTTCCCATATGAAACCAATTTGTGCGCA 570

165 LysfYrHlsleuG1yAlaTyrThrG1yAspAspValArg1leValArgAspAspMetLeu 184
 571 AAATACCCACCTTGGGCGCTACAGGAGAGAGCGCCGATGTCGGTACGACATGCTG 630
 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 631 TGTGCCGGGAAACCCCGGAGGAGCTCATCCAGGGGAGCTCCGGAGGGGCCCTGGTGTGTC 690
 205 LysValAenGlyThrTrrpleuGlnAlaGlyValValSerTrrpGlyGluGlyCysAlaGln 224
 691 AAGGTGAATGGCACCTGGCTGAGGGCGGCGTGCAGCTGGTCCAGGGGAGGGCTGTGCCCG 750
 225 ProAnaArgProGlyYlIeTyrThrArgValThrTyrTyrLeuAspTrrIleHisHsTyr 244
 751 CCACACCGGCGCTGGCATCTACACCGTGTCACTACTTGTGATCTTGGATTCACACATAT 810
 245 ValProLysLysPro 249
 811 GTCCCAAAAAGCCG 825

RESULT 5

US-10-956-157-1043 Application US/10956157
 ; Sequence 1043, Application No. US20050118625A1
 ; Publication No. US20050118625A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William
 ; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 ; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
 ; FILE REFERENCE: 031896-043000 (AM 101081)
 ; CURRENT APPLICATION NUMBER: US/10/956,157
 ; PRIORITY FILING DATE: 2004-10-04
 ; NUMBER OF SEQ ID NOS: 319805
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 1043
 ; LENGTH: 1193
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-956-157-1043

Alignment Scores:
 Pred. No.: 9e-158 Length: 1193
 Score: 1363.00 Matches: 243
 Percent Similarity: 99.188 Conservative: 0
 Best Local Similarity: 99.188 Mismatches: 2
 Query Match: 97.858 Indels: 0
 Gaps: 21
 DB: 21

US-09-598-982c-21 (1-249) x US-10-956-157-1043 (1-1193)

QY 5 lIeValGlyGlnGlnAlaProArgSerLySTrrProTrrpGlnValSerLeuArgVal 24
 DB 143 ATCGTCGGGGGTCAAGAGGAGCCCGCCAGAGCAAGTGGCCCTGGCAGGTGAGCTGAGAGTC 202
 QY 25 HisGlyProTyrTrrpMetHisPheCysGlyYserLeuIleHisProGlnTrrValLeu 44
 DB 203 CAGGGCCCACTGGATGCACTTCTGGCGGGGCTCCCTATCCACCCCAAGGGGGCTG 262
 QY 45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
 DB 263 ACCGCAAGCCACTGGCTGGAGCCGAGCGTCAAGGATCTGGCCCTCGAGGGTGGCAACTG 322
 QY 65 ArgGlyGlnHisLeuTyrTrrpGlnAspGlnLeuLeuProValSerArgGlyIleValHis 84
 DB 323 CCGGAGCAGCAGCTTACTTACAGGAGCCAGCTGCTCCCGTCAAGATCCATCCGTCAC 382
 QY 85 ProGlnPheTyrThrIleGlnIleGlyValAspIleAlaLeuLeuGlnLeuGlnIlePro 104
 DB 383 CCACAGTTCACCCCGCCAGATCGGAGGAGCATGCGCTGTGGAGCTGGAGGAGCCG 442
 QY 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlyIleHisPro 124
 DB 443 GTGAAGGTCTCCAGCGCATCCAGCGTCAAGGATCTGGCCCTCGCAGAGACTTCCCG 502

125 ProGlyMetProCysTrrpValThrGlyTrrpGlyAspValAspAspArgLysArgLeuPro 144
 503 CCGGGAGAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562
 145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
 563 CCGCCATTTCTCTGGAAGCAGGTGAAGGTCCCGCCATTAATGGAAAACCACTTTGTGAGCGCA 622
 165 LysfYrHlsleuG1yAlaTyrThrG1yAspAspValArg1leValArgAspAspMetLeu 184
 623 AAATACCCACCTTGGGCGCTACAGGAGAGCAGCCGCTGTCAGGAGCAGCATGCTG 682
 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 683 TGTGCCGGGAAACCCCGGAGGAGCTCATCCAGGGGAGCTCCGGAGGGGCCCTGGTGTGTC 742
 205 LysValAenGlyThrTrrpleuGlnAlaGlyValValSerTrrpGlyGluGlyCysAlaGln 224
 743 AAGGTGAATGGCACCTGGCTGAGGGCGGCGTGTGCAGCTGGGGGAGGGCTGTGCCCG 802
 225 ProAnaArgProGlyYlIeTyrThrArgValThrTyrTyrLeuAspTrrIleHisHsTyr 244
 803 CCACACCGGCGCTGGCATCTACACCGTGTCACTACTTGTGATCTTGGATTCACACATAT 862
 245 ValProLysLysPro 249
 863 GTCCCAAAAAGCCG 877

RESULT 6

US-10-723-860-6799 Application US/10723860
 ; Sequence 6799, Application No. US20040253606A1
 ; Publication No. US20040253606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Nataasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlobnik, Albert
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
 ; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
 ; FILE REFERENCE: 05882.0193.NPUS01
 ; CURRENT APPLICATION NUMBER: US/10/723,860
 ; PRIORITY FILING DATE: 2003-11-26
 ; CURRENT FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 6799
 ; LENGTH: 1334
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-723-860-6799

Alignment Scores:
 Pred. No.: 1.02e-156 Length: 1334
 Score: 1355.00 Matches: 242
 Percent Similarity: 98.78% Conservative: 0
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 97.27% Indels: 0
 DB: 20

US-09-598-982c-21 (1-249) x US-10-723-860-6799 (1-1334)

QY 5 lIeValGlyGlnGlnAlaProArgSerLySTrrProTrrpGlnValSerLeuArgVal 24
 DB 143 ATCGTCGGGGGTCAAGAGGAGCCCGCCAGAGCAAGTGGCCCTGGCAGGTGAGAGTC 202
 QY 25 HisGlyProTyrTrrpMetHisPheCysGlyYserLeuIleHisProGlnTrrValLeu 44
 DB 203 CCGGAGCAGCAGCTTACTTACAGGAGCCAGCTGCTCCCGTCAAGATCCATCCGTCAC 262
 QY 45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
 DB 263 ACCGCAAGCCACTGGCTGGAGCCGAGCGTCAAGGATCTGGCCCTCGCAGAGACTTCCCG 322

```

QY 65 ArgGIuGlnHisLeuTYrTYrgInaPrgInLleuLeuProValSerArgIleIleValHis 84
Db 323 CGGGACAGACACTTACTACTACAGAGACCAAGCTGCTCGCCGGTGCAGCAAGATCATCTGGCAC 382
QY 85 ProGlnPheTYrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluPro 104
Db 383 CCACAGTTCCTGAAAGCAAGCCCGGAGATCGGACATCGCCCTGTGAGAGTGGAGAGCCGG 442
QY 105 ValIleValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
Db 443 GTGAAGCTCTCCAGCCAGCTCCACAGGGTCAACCCCTGCTCCAGAGACCTTCCCC 502
QY 125 ProGlyMePProCysTrpValThrGlyTPGlyAspValAspAsnAspGluArgLeuPro 144
Db 503 CCGGGATGCTGCTGGGTCACTGGCTGGGGGAGATGTGGCAATGATGAGCGCTCCCA 562
QY 145 ProProPheProLeuIleValIleValValProIleMeGluAsnHisIleCysAspAla 164
Db 563 CCGCCATTTCTCTGAAAGCAAGTGAAGGTCCCATATATGGAACCAATTTGTGACGCA 622
QY 165 LysTYrHisLeuGlyAlaTYrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
Db 623 AANTTACCACTTGGCCCTTACAGGGAGAGCAAGTCCGATTCCTCCGTACAGATGCTG 682
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
Db 683 TGTGGCCGGGAAACCCGGAGGGACTCATGCTCAGGGCGACTCCGGAGGGCCCTGGTGTGC 742
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyIleGlnCysAlaGln 224
Db 743 AAGGTAAATGACCTGGCTGCAAGGGGGGGGTGTGCTGAGGGCCAGAGGGCTGTGCCAG 802
QY 225 ProAsnArgProGlyIleTYrThrArgValThrTYrTYrLeuAspTrpIleHisIleTYr 244
Db 803 CCCAACCGGCTGGCAATACACCCGGTGTACCTTACTTGGACTGGATCCACCACTAT 862
QY 245 ValProIleValPPro 249
Db 863 GTCCCAAAAAGCCG 877

RESULT 7
US-09-954-456-2126
; Sequence 2126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276

```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-2126

Alignment Scores:
Pred. No.: 1,756-155 Length: 1081
Score: 1344.00 Matches: 241
Percent Similarity: 98.37% Conservative: 0
Best Local Similarity: 98.37% Mismatches: 4
Query Match: 96.48% Indels: 0
DB: Gaps: 0

US-09-598-982c-21 (1-249) x US-09-954-456-2126 (1-1081)

QY 5 ILeValIGlyGlyGlnGluAlaProArgSerIleTrpProTProGlnValSerLeuArgVal 24
Db 68 ATCGTTGGGGGTCAAGAGAGCCCGGAGCAAGTGGCCCTCCATCCACCCCAAGTGGTGTCTG 127
QY 25 HisGlyProTYrTrpMetHisPheCysGlyGlySerLeuIleHisPProGlnTrpValLeu 44
Db 128 CCGGACCGATTAATGATGACTTCTGGGGGGCTCCCTCATCCACCCCAAGTGGTGTCTG 187
QY 45 ThrAlaAlaAlaCysValGlyProAspValIleAspLeuAlaAlaLeuArgValGlnLeu 64
Db 188 ACCGAGCGCACTGGCTGGAGACCGGAGCGTCAAGGATCTGGCCCGCCCTCAGGGTGCACCTG 247
QY 65 ArgGIuGlnHisLeuTYrTYrgInaPrgInLleuLeuProValSerArgIleIleValHis 84
Db 248 CCGGAGCGCACTGTTACTTACCAAGGACCAAGCTGCTGGGTCAAGGATCATCTGTGCAC 307
QY 85 ProGlnPheTYrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluPro 104
Db 308 CCACAGTTCCTGAAAGCAAGTGAAGGTCCCATATATGGAACCAATTTGTGACGCA 367
QY 105 ValIleValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
Db 368 GTGAAGGTCTCCAGCCAGCTCCACACCGGTCAACCCCTGGCCCTCCAGAGCCCTTCCCC 427
QY 125 ProGlyMePProCysTrpValThrGlyTPGlyAspValAspAsnAspGluArgLeuPro 144
Db 428 CCGGGAAATGCTGCTGAGGATCACTGGCTGGGGCGATGTGACAAATGATGAGGGCTCCCA 487
QY 145 ProProPheProLeuIleValIleValIleValIleValIleValIleValIleValIle 164
Db 488 CCGCATTTCTCTGAAAGCAAGTGAAGGTCCCATATATGGAACCAATTTGTGACGCA 547
QY 165 LysTYrHisLeuGlyAlaTYrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
Db 548 AANTTACCACTTGGCCCTTACAGGGAGAGCAAGTCCGATTCCTCCAGATGCTG 607
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
Db 608 TGTGGCCGGGAAACCCGGAGGGACTCATGCTCGAGGCTCCCGAGGGCCCTTGGTGTGC 667
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
Db 668 AAGGTGAATGGCACTTGGCTGAGGGGGGGCGGTCACTGGGGGGGAGAGGGCTGTGCCAG 727
QY 225 ProAsnArgProGlyIleTYrThrArgValThrTYrTYrLeuAspTrpIleHisIleTYr 244
Db 728 CCCAACCGGCTGGCAATACACCCGGTGTACCTTACTTGGACTGGATCCACCACTAT 787
QY 245 ValProIleValPPro 249
Db 788 GTCCCAAAAAGCCG 802

RESULT 8
US-09-960-706-680
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1

```

GENERAL INFORMATION:
 APPLICANT: Munger, William E.
 TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
 TITLE OF INVENTION: Gene Expression Profiles
 FILE REFERENCE: 44921-5029-01US
 CURRENT APPLICATION NUMBER: US/09/960,706
 PRIORITY FILING DATE: 2001-09-24
 PRIOR APPLICATION NUMBER: 60/223,323
 PRIOR FILING DATE: 2000-08-07
 PRIOR APPLICATION NUMBER: 09/873,319
 PRIOR FILING DATE: 2001-06-05
 NUMBER OF SEQ ID NOS: 1124
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 680
 LENGTH: 1081
 TYPE: DNA
 ORGANISM: Homo sapiens
 OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
 US-09-960-706-680

Alignment Scores:
 Pred. No.: 1,75e-155 Length: 1081
 Score: 1344.00 Matches: 241
 Percent Similarity: 98.37% Conservative: 0
 Best Local Similarity: 96.37% Mismatches: 4
 Query Match: 96.48% Indels: 0
 Gaps: 0
 DB: 10

US-09-598-982C-21 (1-249) x US-09-960-706-680 (1-1081)
 QY 5 ILEVALIGLYGVLGVALAIPROARSGSERLYSTRPPTROTRGIVALSERLEUARGVAL 24
 DB 68 ATCGTTGGGGGTCAAGAGGCCCCCAGAGCAAGTGGCTTGGCAAGTGGAGTGC 127
 QY 25 HISGLYPROTYRTRPMETHISPHCYSGLYGYSERLEU1EHISPROGLINTRPVALLEU 44
 DB 128 CGGACCCGATCTGATGATGACATTCGCGGGGCTCCCTCATCCACCCCGAGTGGGTGCTG 187
 QY 45 THRVAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAA 64
 DB 188 ACCGAGCCGACCTGCGTGGGAGCCGAGCCTCAAGATCTGGCCCTCAGAGTGGCACTG 247
 QY 65 ARGGLUGLNHISLEU1YTRGIVNAPRGINLEU1SERARG1LE1EVALHIS 84
 DB 248 CGGAGGACGACCTTACTACAGAGCCGCTGCTGAGGATCCATCCGTCGAC 307
 QY 85 PROGINPHEVTRTHRALAGINILEGIVAAAPRILEAIALEU1LEU1LEU1LEU1LEU1LEU1 104
 DB 308 CCAAGTGTCTACCCGCCCAATCCGAGATCCGAGATCCGCTGCTGAGAGGAGCCG 367
 QY 105 VALLYSVALSERSETHISVALHISVTHRVALIHRLEUPROVALSERGIUHRPHEPRO 124
 DB 368 GTGAAGGTCTCCAGCCAGCTCCACAGCTCCAGCTCCGCTCCAGAGACCTTCCCG 427
 QY 125 PROGLYMERPROCYSTRPVALIHRGIVTRGIVYASRVALASRSMRSPGLUARGLEUPRO 144
 DB 428 CCGGGATCCCGTGGCTGCTCACTGGCTGGGAGATGAGAAATATAGAGGCCCTCCCA 487
 QY 145 PROPROPHEPROLEU1YSGINVALLYSVALIPROILEMERTGLVASHIS1IECYSPASPA 164
 DB 488 CCGCCATTTCTCTGTAAGAGGTGGTGAAGTCCCAATATGAAACCAKATTTGTGACGA 547
 QY 165 IYSTRYHISLEU1GIVAAIYTRGIVYASRVALIARGILEVALARGHAPSPHECTLEU 184
 DB 548 AAATATACACTTGGGGCTTACACGGGAGAGCAAGCTCCGATCGCTGAGACGACATGCTG 607
 QY 185 CYGALAGLYASRTHRARGARGASRSEYCYGGINGLYASRSEGLIYTRPVALCYG 204
 DB 608 TGTGCGGGAGAACCCGAGGAGCTCATCCAGGGGACCTCCGAGAGGCCCTGCTGTC 667
 QY 205 IYVVALASRGLYTRHTRPHEU1NIAAGIYVALVASETRTRGIVGIVYCSVALAGIN 224

DB 668 AAGGTAAATGGCACCTGAGCCGAGCCGGCTGAGCTGGGGGAGAGGCTGTGCCAG 727
 QY 225 PROANRARGPROGLIYIETRYTRHARGVALIHRLEU1AEPTRP1LEHIS1TYR 244
 DB 728 CCAACCGGCTTGGAGATCTACACCCGTTGATCCTACTTGGAGCTGATCCACACTAT 787
 QY 245 VALPROLYSILYSPRO 249
 DB 788 GTCCCAAAAAGCCG 802

RESULT 9
 US-09-873-319-427
 Sequence 427, Application US/09873319A
 Publication No. US20030134324A1
 GENERAL INFORMATION:
 APPLICANT: Munger, William E.
 APPLICANT: Kulkarni, Prakash
 APPLICANT: Getzenberg, Robert H.
 APPLICANT: Waga, Iwao
 APPLICANT: Yamamoto, Jun
 TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
 TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
 FILE REFERENCE: 44921-5029-US
 CURRENT APPLICATION NUMBER: US/09/873,319A
 PRIORITY FILING DATE: 2001-06-05
 EARLIER APPLICATION NUMBER: US 60/223,323
 PRIOR FILING DATE: 2000-08-07
 NUMBER OF SEQ ID NOS: 755
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 427
 LENGTH: 1081
 TYPE: DNA
 ORGANISM: Homo sapiens
 OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
 US-09-873-319-427

Alignment Scores:
 Pred. No.: 1,75e-155 Length: 1081
 Score: 1344.00 Matches: 241
 Percent Similarity: 98.37% Conservative: 0
 Best Local Similarity: 96.48% Mismatches: 4
 Query Match: 96.48% Indels: 0
 Gaps: 0
 DB: 10

US-09-598-982C-21 (1-249) x US-09-873-319-427 (1-1081)
 QY 5 ILEVALIGLYGVLGVALAIPROARSGSERLYSTRPPTROTRGIVALSERLEUARGVAL 24
 DB 68 ATCGTTGGGGGTCAAGAGGCCCCCAGAGCAAGTGGCTTGGCAAGTGGAGTGC 127
 QY 25 HISGLYPROTYRTRPMETHISPHCYSGLYGYSERLEU1EHISPROGLINTRPVALLEU 44
 DB 128 CGGACCCGATCTGATGATGACATTCGCGGGGCTCCCTCATCCACCCCGAGTGGGTGCTG 187
 QY 45 THRVAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAAIAA 64
 DB 188 ACCGAGCCGACCTGCGTGGGAGCCGAGCCTCAAGATCTGGCCCTCAGAGTGGCACTG 247
 QY 65 ARGGLUGLNHISLEU1YTRGIVNAPRGINLEU1SERARG1LE1EVALHIS 84
 DB 248 CGGAGGACGACCTTACTACAGAGCCGCTGCTGAGGATCCATCCGTCGAC 307
 QY 85 PROGINPHEVTRTHRALAGINILEGIVAAAPRILEAIALEU1LEU1LEU1LEU1LEU1LEU1 104
 DB 308 CCAAGTGTCTACCCGCCCAATCCGAGATCCGAGATCCGCTGCTGAGAGGAGCCG 367
 QY 105 VALLYSVALSERSETHISVALHISVTHRVALIHRLEUPROVALSERGIUHRPHEPRO 124
 DB 368 GTGAAGGTCTCCAGCCAGCTCCACAGCTCCAGCTCCGCTCCAGAGACCTTCCCG 427
 QY 125 PROGLYMERPROCYSTRPVALIHRGIVTRGIVYASRVALASRSMRSPGLUARGLEUPRO 144

```

DB 428 CCGGGATGCGGTGCTGCATCCTGGCGGGCGGATGTGTGACATATGATGAGCGCCCTCCCA 487
QY 145 ProProPheProLeuLysGlnValLysValProIleMetGlnuAhnIleCysAspAla 164
DB 488 CCGCCATTTTCTCTGAAAGCAAGGTAAGGTCCCATATGGAANAACAATTTTGTACCGCA 547
QY 165 LysTyrHisLeuGlnValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 548 AAATACCCACCTTGGCCCTTACACGGGAGACGAGTCCGCACTCCGTCCGTGACGACATGCTG 607
QY 185 CysAlaGlyAsnThrArgAspAspSerCysGlnGlyAspSerGlyValProLeuValCys 204
DB 608 TGTGCGGGAGAACCCGGAGAGACTATGCTCCAGAGGAGACTCCGGAGGGCCCTGGGTGTCG 667
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValIleValSerTrpGlyGluGlyCysAlaGln 224
DB 668 AAGTGAATGGCACTGGCTGGCGAGGGGGGTGTGTGAGCTGGGGCCGAGGGGCTGTGCCAG 727
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrIleuAspTrpIleHisHisTyr 244
DB 728 CCCAACCGGCTGTGACCTTACACCCCGTGTACCTACTGATGGATCCACCACTAT 787
QY 245 ValProLysLysPro 249
DB 788 GTCCCAAAAAGCCG 802

```

RESULT 10

```

US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-155

```

```

Alignment Scores:
Pred. No.: 1,75e-155 Length: 1081
Score: 1344.00 Matches: 241
Percent Similarity: 98.37% Conservative: 0
Best Local Similarity: 98.37% Mismatches: 4
Query Match: 96.48% Indels: 0
DB: 10 Gaps: 0

```

```

US-09-598-982c-21 (1-249) x US-09-873-367C-155 (1-1081)
QY 5 ILeValGlyGlnGlnAlaProAspSerLysTrpProTrpGlnValSerLeuArgVal 24
DB 68 ATCGTTGGGGGCTCAGAGGCGCCCAAGAGCAAGTGGCCCTGGCGAGGTGAGCCTGAGAGTTC 127
QY 25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44

```

```

DB 128 CCGACCGATCTGATGCATCCTTCTGCGGGGGCTCCATCCACCCCAAGTGGTCTG 187
QY 45 ThrAlaAlaCysValGlyProAspValLysAspLeuAlaLeuArgValGlnLeu 64
DB 188 ACCGAGCGCCACTGGGTGGACCCGAGCGTCAAGATCTGGCCCTCAGGGTGCACCTG 247
QY 65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
DB 248 CCGGAGCAGCACCTCTACTACAGGACCGAGTCTGGCCGGTCAAGAGATCATCTGTGAC 307
QY 85 ProGlnPheTyrThrAlaGlnIleGlyValAspIleAlaLeuLeuGlnLeuGlnLeu 104
DB 308 CCACAGTTCTTACACCGCCAGATCGAGCGGACATCGCCCTGTGACTGAGAGAGCCG 367
QY 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
DB 368 GTGAAAGTCTCCAGCCACGTCACACCGTCACTGCCCTGGCTTCAAGACCTTCCCC 427
QY 125 ProGlyMetProCysTrpValThrGlyTyrGlyAspValAspAsnAspGluArgLeuPro 144
DB 428 CCGGGGATGCGGTGGGTGCTGCTGGGGTCACTGGGTGGGGCAATGTGACATGATGAGCGCTCCCA 487
QY 145 ProProPheProLeuLysGlnValLysValProIleMetGlnuAhnIleCysAspAla 164
DB 488 CCGCCATTTTCTCTGAAAGCAAGGTAAGGTCCCATATGGAANAACAATTTTGTACCGCA 547
QY 165 LysTyrHisLeuGlnValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 548 AAATACCCACCTTGGCCCTTACACGGGAGACTATGCTCCAGAGGAGACTCCGGAGGGCCCTGGGTGTCG 607
QY 185 CysAlaGlyAsnThrArgAspAspSerCysGlnGlyAspSerGlyValProLeuValCys 204
DB 608 TGTGCGGGAGAACCCGGAGAGACTATGCTCCAGAGGAGACTCCGGAGGGCCCTGGGTGTCG 667
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValIleValSerTrpGlyGluGlyCysAlaGln 224
DB 668 AAGTGAATGGCACTGGCTGGCGAGGGGGGTGTGTGAGCTGGGGCCGAGGGGCTGTGCCAG 727
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrIleuAspTrpIleHisHisTyr 244
DB 728 CCCAACCGGCTGTGACCTTACACCCCGTGTACCTACTGATGGATCCACCACTAT 787
QY 245 ValProLysLysPro 249
DB 788 GTCCCAAAAAGCCG 802

```

RESULT 11

```

US-09-873-367C-714
; Sequence 714, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0

```

```

: SEQ ID NO 714
: LENGTH: 1081
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-873-367C-714
Alignment Scores:
Alignment No.: 1,75e-155      Length: 1081
Score: 1344.00             Matches: 241
Percent Similarity: 98.37%   Conservative: 0
Best Local Similarity: 98.37% Mismatches: 4
Query Match: 96.48%         Indels: 0
DB: 10                      Gaps: 0

US-09-598-982c-21 (1-249) x US-09-873-367C-714 (1-1081)
OY 5 11eva1g1yglng1u1a1p1roa1rse1y1t1p1r1g1n1va1se1ru1a1r1g1a1
DB 68 ATCGTTGGGGGTCAGAGAGCCGCCCAAGACAAAGTGGCCCTGGCAGGTGAGTGC
OY 25 H18g1P1ro1y1r1t1p1m1e1h1s1p1h1e1c1y1e1g1y1s1e1r1e1u1i1e1h1s1p1r1o1g1n1t1r1p1a1l1e1u
DB 128 CGGACCGGATCTGATGCACTTCTGGCGGGGCTCCCTCATCCACCCCGAGTGGGTGGT
OY 45 Thr1aa1a1a1Cyeva1g1y1p1roa1rse1y1a1l1y1a1s1p1e1u1a1a1a1e1u1a1r1g1n1e1u
DB 188 ACCGACGGCCACTGGTGGGAGCCGACCTGCAAGATCTGGCCCTCGAGGTTGCACTG
OY 65 Ar1g1u1g1n1h1s1e1u1y1r1y1g1n1s1p1n1e1u1e1p1r1o1a1s1e1r1a1g1l1e1l1e1v1a1h1s
DB 248 CGGGAGACAGCACTGACTACAGAGGACAGGCTGCTGGCCGATGCAATGATGAGGCCCTCCA
OY 85 Pr1o1g1n1p1h1e1y1r1t1r1a1a1g1n1l1e1g1y1a1a1s1p1i1e1a1l1e1u1e1g1u1e1n1g1u1p1r1o
DB 308 CCACAGTTCCTACACCGCCAGATCGGAGCGGACATCGCCCTCGAGCTGAGGAGCGCG
OY 105 Val1y1v1a1s1e1r1s1e1h1s1v1a1h1s1t1h1r1v1a1l1t1h1r1e1u1p1r1o1a1s1e1r1g1u1t1h1r1p1h1e1r1o
DB 368 GTGAAGTTCCTCAAGCCAGCTCAACAGTCACTCCCTGGCTCGAGACCTTTCCCC
OY 125 Pr1o1g1m1e1t1p1r1o1c1y1e1t1r1p1v1a1t1h1r1g1y1t1r1p1g1y1a1s1p1v1a1s1p1n1s1p1g1u1a1r1g1e1u1p1r1o
DB 428 CGGGGGATCCCGTGGTGGTCACTGGCTGGGGGATGATGAGCAATGATGAGGCCCTCCA
OY 145 Pr1o1p1r1o1p1h1e1r1o1e1u1l1y1e1g1n1v1a1l1y1s1v1a1l1p1r1o1l1e1m1e1r1g1u1a1n1h1s1i1e1c1y1a1s1p1a1a
DB 488 CGCCCAATTCCTCTGTAAGGAGGTAAGGTCCCAATGATGAAACCAACCACTTTGTCAGCGCA
OY 165 Lys1y1r1h1s1e1u1g1y1a1a1t1y1t1h1r1g1y1a1s1p1a1r1v1a1a1r1g1l1e1v1a1a1r1g1a1s1p1a1r1p1m1e1l1e1u
DB 548 AATATCACCTTGGCCCTACACGGGAGGAGGACAGCTCCGATGCTCCGTAACAGACATGCTG
OY 185 Cy1a1a1g1y1a1s1p1n1t1h1r1g1a1r1g1a1s1e1r1c1y1s1g1n1g1y1a1s1p1s1e1r1g1y1p1r1o1e1u1v1a1c1y1s
DB 608 TGTGGCCGGGAAACCCCGAGGGGACTATCCAGGGGGAATCCGGAGGGGCCCTGGGTGGTGC
OY 205 Lys1y1v1a1s1e1r1g1y1t1r1p1e1u1g1n1a1g1y1v1a1l1v1a1s1e1r1t1p1g1y1g1u1g1y1a1a1g1n
DB 668 AAGGTGAATAGGCACTCGCTGGCAGGGGGGGGCTGTGATCAGCTGGGGGGCGAGGGGCTGTGGCCCG
OY 225 Pr1o1a1n1a1r1g1y1r1e1t1y1t1h1r1a1a1g1n1l1e1g1y1a1a1s1p1i1e1a1l1e1u1e1g1u1e1n1g1u1p1r1o
DB 728 CCDAACGGGCTGGCACTCAACCGGTGACCTTACTTGGACTGGATCCACACTAT
OY 245 Val1p1r1o1y1a1l1y1e1p1r1o
DB 788 GTCCCCAAAGGCGG 802

```

```

: APPLICANT: Avalon Pharmaceuticals, Inc.
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
: FILE REFERENCE: 689290-189
: CURRENT FILING DATE: 2004-05-12
: PRIORITY FILING DATE: 2004-05-12
: PRIORITY FILING DATE: US/09/873,367
: PRIORITY FILING DATE: 2001-06-05
: PRIORITY FILING DATE: US/09/954,531
: PRIORITY FILING DATE: 2001-09-18
: PRIORITY FILING DATE: US/09/954,456
: PRIORITY FILING DATE: 2001-09-25
: PRIORITY FILING DATE: US/09/962,436
: PRIORITY FILING DATE: 2001-09-25
: PRIORITY FILING DATE: US/09/962,832
: PRIORITY FILING DATE: 2001-09-25
: PRIORITY FILING DATE: US/09/964,824
: PRIORITY FILING DATE: 2001-09-27
: PRIORITY FILING DATE: US/09/967,768
: PRIORITY FILING DATE: 2001-09-28
: PRIORITY FILING DATE: US/09/968,007
: PRIORITY FILING DATE: 2001-10-02
: PRIORITY FILING DATE: US/09/969,347
: PRIORITY FILING DATE: 2001-10-02
: PRIORITY FILING DATE: US/09/969,708
: PRIORITY FILING DATE: 2001-10-03
: REMAINING PRIORITY APPLICATION DATA REMOVED - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 8447
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 155
: LENGTH: 1081
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-843-641A-155
Alignment Scores:
Alignment No.: 1,75e-155      Length: 1081
Score: 1344.00             Matches: 241
Percent Similarity: 98.37%   Conservative: 0
Best Local Similarity: 98.37% Mismatches: 4
Query Match: 96.48%         Indels: 0
DB: 21                      Gaps: 0

US-09-598-982c-21 (1-249) x US-10-843-641A-155 (1-1081)
OY 5 11eva1g1yglng1u1a1p1roa1rse1y1t1p1r1g1n1va1se1ru1a1r1g1a1
DB 68 ATCGTTGGGGGTCAGAGAGCCGCCCAAGACAAAGTGGCCCTGGCAGGTGAGTGC
OY 25 H18g1P1ro1y1r1t1p1m1e1h1s1p1h1e1c1y1e1g1y1s1e1r1e1u1i1e1h1s1p1r1o1g1n1t1r1p1a1l1e1u
DB 128 CGGACCGGATCTGATGCACTTCTGGCGGGGCTCCCTCATCCACCCCGAGTGGGTGGT
OY 45 Thr1aa1a1a1Cyeva1g1y1p1roa1rse1y1a1l1y1a1s1p1e1u1a1a1a1e1u1a1a1e1u1a1r1g1n1e1u
DB 188 ACCGACGGCCACTGGTGGGAGCCGACCTGCAAGATCTGGCCCTCGAGGTTGCACTG
OY 65 Ar1g1u1g1n1h1s1e1u1y1r1y1g1n1s1p1n1e1u1e1p1r1o1a1s1e1r1a1g1l1e1l1e1v1a1h1s
DB 248 CGGGAGACAGCACTGACTACAGAGGACAGGCTGCTGGCCGATGCAATGATGAGGCCCTCCA
OY 85 Pr1o1g1n1p1h1e1y1r1t1r1a1a1g1n1l1e1g1y1a1a1s1p1i1e1a1l1e1u1e1g1u1e1n1g1u1p1r1o
DB 308 CCACAGTTCCTACACCGCCAGATCGGAGCGGACATCGCCCTCGAGCTGAGGAGCGCG
OY 105 Val1y1v1a1s1e1r1s1e1h1s1v1a1h1s1t1h1r1v1a1l1t1h1r1e1u1p1r1o1a1s1e1r1g1u1t1h1r1p1h1e1r1o
DB 368 GTGAAGTTCCTCAAGCCAGCTCAACAGTCACTCCCTGGCTCGAGACCTTTCCCC
OY 125 Pr1o1g1m1e1t1p1r1o1c1y1e1t1r1p1v1a1t1h1r1g1y1t1r1p1g1y1a1s1p1v1a1s1p1n1s1p1g1u1a1r1g1e1u1p1r1o
DB 428 CGGGGGATCCCGTGGTGGTCACTGGCTGGGGGATGATGAGCAATGATGAGGCCCTCCA

```

QY 145 ProPrpPheProLeuLySgInValLySValProIleMetGluAAsnHisIleCysAspAla 164
 Db 488 CCCCAATTCCTCTGAAGCGGTGAAGGTCCTCCCATTAATGGAAAAACAACATTTGTGACGCA 547
 QY 165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
 Db 548 AAATACCACTTGTGGCCCTTACCGGAGACGACGTCCTCCGATCCGTGACGACATGCTG 607
 QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 Db 608 TGTGCCGGGAAACACCCGGAGGAGCACTATGCGAGGCGACTCCGGAGGGCCCTGGTGTGC 667
 QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlyGlyCysAlaGln 224
 Db 668 AAGGTAAATGGACCTGGCTGGACGGCGGGGTGTGATGCTGAGGGCCAGTGTGCCAG 727
 QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIstYr 244
 Db 728 CCCAACCGGCTTGGCATCTACACCCGTTGTCTACTACTTGGACTGGATCCACCACTAT 787
 QY 245 ValProLySlySPro 249
 Db 788 GTCCCAAAAAGCCG 802

RESULT 13

US-10-843-641A-714
 ; Sequence 714, Application US/10843641A
 ; Publication No. US20050064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE REFERENCE: Signature Gene Sets
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; PRIOR FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US/09/967,768
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/968,007
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,347
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,708
 ; PRIOR FILING DATE: 2001-10-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 8447
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 714
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-843-641A-714

Alignment Scores:

Pred. No.: 1,75e-155 Length: 1081
 Score: 1344.00 Matches: 241
 Percent Similarity: 98.37% Conservative: 0
 Best Local Similarity: 98.37% Mismatches: 4
 Query Match: 96.48% Indels: 0
 DB: 21 Gaps: 0

US-09-598-982C-21 (1-249) x US-10-843-641A-714 (1-1081)

QY 5 IleValGlyGlyGlnAlaProArgSerIlyTrpProTrpGlnValSerIleuArgVal 24
 Db 68 ATCGTTGGGGGTCAAGAGCCCCCAAGAGCAAGTGGCCCTGGCAAGTGGCTGGAGAGTC 127
 QY 25 HisGlyProTyrTrpMetHisPheCysGlyGlySerIleuIleHisProGlyThrTrpValIleu 44
 Db 128 CCGGACCGATCTGGATGATGATCTTGTCCGGGGCTCCATCCACCCCAAGTGGGTGGT 187
 QY 45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaIleuArgValGlnIleu 64
 Db 188 ACCGGAGCGCACTGGGTGGAGCCGAGACGTCMAAGATCTGGCCCTCAGGGTGGCAACTG 247
 QY 65 ArgGlyGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
 Db 248 CCGGAGCGACACTTACTACCAAGACCAAGCTGCTGGCGGCTCAGAGGATCATCTGTGAC 307
 QY 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaIleuLeuGlyIleuGlyIleuPro 104
 Db 308 CCACAGTCTTACACCGCCAGATCGGAGGACATCGCTGTGGAGCTGGAGGACCG 367
 QY 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlyIleuTrpPhePro 124
 Db 368 GTGAAGGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 427
 QY 125 ProGlyMetProCysTrpValThrGlyTyrGlyAspValAspAspAspGlyIleuGlyIleuPro 144
 Db 428 CCGGGGATGCTCGTGGGTCACTGGCTGGGGCGATGTGGACAAATGATGAGCGCCCTCCCA 487
 QY 145 ProPrpPheProLeuLySgInValLySValProIleMetGluAAsnHisIleCysAspAla 164
 Db 488 CCCCAATTCCTCTGAAGCGGTGAAGGTCCTCCCATTAATGGAAAAACAACATTTGTGACGCA 547
 QY 165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
 Db 548 AAATACCACTTGTGGCCCTTACCGGAGACGACGTCCTCCGATCCGTGACGACATGCTG 607
 QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 Db 608 TGTGCCGGGAAACACCCGGAGGAGCACTATGCGAGGCGACTCCGGAGGGCCCTGGTGTGC 667
 QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlyGlyCysAlaGln 224
 Db 668 AAGGTAAATGGACCTGGCTGGACGGCGGGGTGTGATGCTGAGGGCCAGTGTGCCAG 727
 QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIstYr 244
 Db 728 CCCAACCGGCTTGGCATCTACACCCGTTGTCTACTACTTGGACTGGATCCACCACTAT 787
 QY 245 ValProLySlySPro 249
 Db 788 GTCCCAAAAAGCCG 802

RESULT 14

US-10-843-641A-5153
 ; Sequence 5153, Application US/10843641A
 ; Publication No. US20050064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE REFERENCE: Signature Gene Sets
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; PRIOR FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832

PRIOR FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US/09/964,824
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: US/09/967,768
 PRIOR FILING DATE: 2001-09-28
 PRIOR APPLICATION NUMBER: US/09/968,007
 PRIOR FILING DATE: 2001-10-02
 PRIOR APPLICATION NUMBER: US/09/969,347
 PRIOR FILING DATE: 2001-10-02
 PRIOR APPLICATION NUMBER: US/09/969,708
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 8447
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 5153
 LENGTH: 1081
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-843-641A-5153

Alignment Scores:
 Pred. No.: 175e-155 Length: 1081
 Score: 1344.00 Matches: 241
 Percent Similarity: 98.37% Conservative: 0
 Best Local Similarity: 96.48% Mismatches: 4
 Query Match: 21 Indels: 0
 DB: Gaps: 0

US-09-598-982c-21 (1-249) x US-10-843-641A-5153 (1-1081)

5 lIeValGlyGlnGlnAlaProArGserLySTrPrOTrGlnValSerLeuArGVal 24
 68 ATCGTTGGGGGTCAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGGAGAGTCC 127
 25 HisGlyProTyTrTmMethHisPheCysGlyGlySerLeuIleHisProGlnTrpValIleu 44
 128 CGCGACCGATATGGATGTGACTTTCGGGGGGCTCCCTCATCCACCCCGAGTGGGTGG 187
 45 ThrAlaAlaAlaCysValGlyProArPValLySArPLeuAlaAlaLeuArGValGlnLeu 64
 188 ACCGAGCCGACCTGGCGTGGAGCCGGACGTCAGAGATCTGGCCCTCAGGGTGAACCTG 247
 65 ArGInGlnHisLeuTyTrTyrGlnAAsPglNleuLeuProValSerArGtlleIleValHis 84
 248 CGGGAGCCAGACCTCTACTTACCAAGACCGACTGGCCCGGTACGAGCATTCGTTGAC 307
 85 ProGlnPheTyTrThraAlaGlnIleGlyAlaAAsPrlleAlaLeuGlnLeuGlnGlnPro 104
 308 CCACAGTTCCTTACACCGCCAGATCCGAGGAGGACATCCGCTGCTGGAGTGGAGAGCCG 367
 105 ValLySValSerSerHisValHisThraValThrLeuProProlASerGluThrPhePro 124
 368 GTGAAGGTCCTCCAGCCACCTCCACCGTCCACCCCTCCCGCCCTCAGAGACCTTCCCC 427
 125 ProGlnMetProCysTrpValThrGlyTrpGlyAAsPValAAsPAsnAsPglNArGlnLeuPro 144
 428 CCGGGAGTCCCGTGGGTGCTGGGTGCTGGGGGAGTGGAGTGGAGTGGAGTGGAGTGG 487
 145 ProProPheProLeuLySValIleValLySValProlIleMetGlnAAsHisIleCysArPAla 164
 488 CGGCCAATTTCTCTAAGAGGAGTGAAGGTCCTCCCATATATGAAACCAACATTTTGTAGCG 547
 165 LySlyrThIleLeuGlyAlaIleTyTrThrGlyAAsPArPValArGtlleValArGArPArPMedLeu 184
 548 AAATGTCACACCTTGGCCCTTACACGGGAGACGAGTCCCGCATGCTCCGTACCAACATGGCTG 607
 185 CysAlaIleAAsnThraArGArPArPArPArPArPArPArPArPArPArPArPArPArPArP 204
 608 TGTGGCCGAAACACCCCGAGGAGTCTATATCCAGAGGAGTCCCGAGAGGAGCCCTGGGTGTC 667
 205 LySValAAsnGlyThrTrpLeuGlnAlaGlyValIleSerTrpGlyGlnGlyCysAlaGln 224
 668 AAGGTGAATGGACCTGGTGGAGGGCGTGGTCACTGGGGGAGGAGGAGGAGGAGGAGGAGG 727

225 ProknaArgProGlyIleTyTrThraArGValThrTyTrIleuAAsPrlleHisIleTyTr 244
 728 CCGAACCCTGGCATCTTACACCGGTCACCTTACTTGGACTGGATCCACCACTAT 787
 245 ValProLySlySPro 249
 788 GTCCCAAAAAGCCG 802

RESULT 15
 US-10-275-505-27
 ; Sequence 27, Application US/10275505
 ; Publication No. US20040081961A1
 GENERAL INFORMATION:
 APPLICANT: INCYTE GENOMICS, INC.
 APPLICANT: DELSEANE, Angelo M.; LAL, Preeti G.
 APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra
 APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
 APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
 APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
 APPLICANT: AZIZZAI, Yalda; ELIOTT, Vicki S.
 APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
 APPLICANT: YANG, Junming; HERNANDEZ, Roberto
 APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
 APPLICANT: REDDY, Roopa M.; YUE, Henry
 APPLICANT: TANG, Y. Tom
 TITLE OF INVENTION: PROTEASES
 FILE REFERENCE: PI-0085 USN
 CURRENT APPLICATION NUMBER: US/10/275,505
 CURRENT FILING DATE: 2002-11-04
 PRIOR APPLICATION NUMBER: PCT/US01/14651
 PRIOR FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/209,402
 PRIOR FILING DATE: 2000-06-01
 PRIOR APPLICATION NUMBER: 60/207,477
 PRIOR FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: 60/205,803
 PRIOR FILING DATE: 2000-05-17
 PRIOR APPLICATION NUMBER: 60/203,566
 PRIOR FILING DATE: 2000-05-11
 PRIOR APPLICATION NUMBER: 60/202,082
 PRIOR FILING DATE: 2000-05-04
 NUMBER OF SEQ ID NOS: 28
 SOFTWARE: PERL Program
 SEQ ID NO 27
 LENGTH: 2662
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURES:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No: 7472460CB1
 US-10-275-505-27

Alignment Scores:
 Pred. No.: 4.13e-153 Length: 2662
 Score: 1329.00 Matches: 238
 Percent Similarity: 97.14% Conservative: 0
 Best Local Similarity: 97.14% Mismatches: 7
 Query Match: 95.41% Indels: 0
 DB: Gaps: 0

US-09-598-982c-21 (1-249) x US-10-275-505-27 (1-2662)

5 lIeValGlyGlnGlnAlaProArGserLySTrPrOTrGlnValSerLeuArGVal 24
 701 ATCGTTGGGGGTCAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGGAGAGTCC 760
 25 HisGlyProTyTrTmMethHisPheCysGlyGlySerLeuIleHisProGlnTrpValIleu 44
 761 CGCGACCGATATGGATGTGACTTTCGGGGGGCTCCCTCATCCACCCCGAGTGGGTGG 820
 45 ThrAlaAlaAlaCysValGlyProArPValLySArPLeuAlaAlaLeuArGValGlnLeu 64

```

Db      821 ACCGACCGCAGCTGCGTGGGACCGGACGTCAGAGGATCTGGCCCGCTCAGGGTGCACACTG 880
QY      65  ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
Db      881 CCGGAGCAGACCTCTACTACCGAGGACCGAGCTGTGCCGGGTGACGAGATCCGTGCAC 940
QY      85  ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluPro 104
Db      941 CCACAGTTCTACACCCCGCCAGATCGGAGGAGACATCGCCCTGTGAGCTGGAGGCCG 1000
QY      105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnPhePhePro 124
Db      1001 GTGAAAGTCTCCAGCCAGCTCCACACCGTCACTCCCTGCCCTCAGAGACTTCCCC 1060
QY      125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
Db      1061 CCGGGGATGCGCGTGGGTCACTGGCTGGGGGATGTGGACATGATGAGGGCCCTCCCA 1120
QY      145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
Db      1121 CCGCCATTTCCCTGAAAGCATGTGAAGGTCCCCATATGGAAMAAACCATTTGTGACGCA 1180
QY      165 LysTyrHisLeuGluValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
Db      1181 AAATACCACCTTGGCCGCTACACCGGAGAGACGATCCGCATCGTCCGTGACACATGCTG 1240
QY      185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
Db      1241 TGTGCCGGGAAACCCCGGAGGGACTCATGCGCAGGGCGGACTCCGGAGGGCCCTGGGTGTC 1300
QY      205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
Db      1301 AAGGTGAATGGCACCTGGCTGACGGCGGGCGGTGCAGATGGGGAGAGGGCTGTGCCAG 1360
QY      225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIleTyr 244
Db      1361 CCCAACCGGCTGTGGCATCTACACCCCGTGTCACTACTTGGACTGGAATCCACCACTAT 1420
QY      245 ValProLysLysPro 249
Db      1421 GTCCCCAAAAAGCCG 1435

```

Search completed: August 27, 2005, 23:11:52
 Job time : 613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 19:35:09 ; Search time 178 Seconds
(without alignments)
2288.949 Million cell updates/sec

Title: US-09-598-982C-21
Perfect score: 1393
Sequence: 1 LBRKRVGGQEARSRKMPWQV.....ITTRVTVYLDIMTHVYRPPK 249

Scoring table:
BLOSSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delpop 6.0 , Deltext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ _p2n.model -DEV=xlip
-O=/cgn2_1/spool_p/US09598982/rnuc 25082005 165518 269/ app_query.fasta_1.391
-DB=Issued Patente NA -QFMT=fastap -SUPFIX=rxni -MINMATCH=0.1 -LOOFCFL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=humann40.cdi
-LIST=45 -DCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982@cgn1_1.105 @runat 25082005 165518 269 -NCPU=6 -ICPU=3
-NO WMAP -LANG=ENGLISH -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patente_NA:*\n1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*\n2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*\n3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*\n4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*\n5: /cgn2_6/prodata/1/ina/ECTUS_COMB.seq:*\n6: /cgn2_6/prodata/1/ina/backfills1.seq:*

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 12 rows of sequence alignment data.

Table with columns: 13, 1091, 78.3, 1108, 2, US-08-978-404B-20, Sequence 20, Appli. Contains alignment details for sequence 20.

RESULT 1
US-09-079-970A-4
Sequence 4, Application US/09079970A
Patent No. 6274366
GENERAL INFORMATION:
APPLICANT: Mafilt, Mark A.
APPLICANT: Miles, Andrew L.
APPLICANT: Haak-Frendscho, Mary
TITLE OF INVENTION: Enzymatically-Active Recombinant Human
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSER: Intellectual Property Department
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WISCONSIN
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079, 970A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leone, Joseph T.
REGISTRATION NUMBER: 37, 170
REFERENCE/DOCKET NUMBER: 34506, 073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:

```

?       LENGTH: 771 base pairs
?       TYPE: nucleic acid
?       STRANDEDNESS: single
?       TOPOLOGY: linear
?       MOLECULE TYPE: DNA (genomic)
?       HYPOTHETICAL: NO
?       ANTI-SENSE: NO
?       ORIGINAL SOURCE:
?       ORGANISM: Homo sapiens
?       FEATURE:
?       NAME/KEY: CDS
?       LOCATION: 7..753
?       FEATURE:
?       NAME/KEY: misc_signal
?       LOCATION: 7..18
?       US-09-079-970A-4

```

```

Alignment Scores:
Pred. No.: 3.11e-150          Length: 771
Score: 1387.00              Matches: 248
Percent Similarity: 99.60%   Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 99.57%         Indels: 0
DB: 3                         Gaps: 0

```

US-09-598-982c-21 (1-249) x US-09-079-970A-4 (1-771)

```

QY 1 LeuGluYlYsRgIleValGIyGIInGUaLaPProArGserIyTPrPProTgInVal 20
DB 7 CTGGAAAGAAAGATCGTCGGGGGTCAAGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGTIG 66
QY 21 SerLeuArGValHISGLYProTYTTPMethHISpHeCySGIyGIYserLeuIleHISPro 40
DB 67 AGCCTGAGGTCCACCGCCCACTGAGTCACTTCGCGGGGGCTCCCTCATCCACCCC 126
QY 41 GluTrpValleuThrAlaAlaAaCyValGIYPrAspValIlyserIleuAlaAlaLeu 60
DB 127 CATGGGTGGTCAGCCGAGGCACTGCTGGGACCCGAGAGTCAAGGATCGCCGCGCTC 186
QY 61 ArgValGlnleuArGInGInHISleuTYTYTGINAspGlnleuLeuPProValSerArG 80
DB 187 AGGTGCACACTGGGGAGGAGCACTCTTCAACAGACCAAGTGTGCGCCGTGAGCAAG 246
QY 81 IleIleValHISProGlnPheTYTThrAlaGlnIleGIyValAspIleAlaLeuLeuGlu 100
DB 247 ATCATGCTGACCCCAAGTCTTCAACCGCCGAGATGGAGCGGACATCCCTGCTGGAG 306
QY 101 LeuGluGluProValIlyserIleHISThrValThreuPProAlaSer 120
DB 307 CTGGAGAGCCGGTGAAGGTCTCAACCCAGTCAACAGTCAACCTGCCCTGCTGCTCA 366
QY 121 GIuThrPheProGlnYmECProCysTrpValThrGIYTrpGIYAspValAspAsp 140
DB 367 GAGACCTTCCCGCGGGGANTGCCGTGCTGCTGCACTGGGGCGAATGTGACAATAT 426
QY 141 GIuArGLeuPProPProPheProLeuIyGIInValIlyValProIleuGIuAsnHIS 160
DB 427 GAGCGCTCCCAACCGCATTCTCTGAAGCAAGTGAAGTCCCAATATGAAAAACAC 486
QY 161 IleCyAspAlaIlyserIleHISleuGIYAlaTYTThrGIYAspAspValAlaGIleValArG 180
DB 487 ATTGTGACGCAAAATACACTTGGCGCTACACGGGAGACGACCTCCGCAATCCGCT 546
QY 181 AspAspMetLeuCYeAlaGIYAsnThrArGAspSerCyGIInGIYAspserGIYGIY 200
DB 547 GACGATCATGCTGTGTCCGGGAACAACCCGAGGAGGATCACTGCAAGGCGCACTCCGAGGG 606
QY 201 ProLeuValCYeAlYsValAsnGIYThrTrpLeuGlnAlaGIYValIlyValserTrpGIYGIY 220
DB 607 CCCCTGGGTGGCAAGGTGAATGGCACTGCTGCAAGGCGGGCGTGGTCACTGGGGCGAG 666
QY 221 GIYCyAlaGIYProAspArPProGIYIleTYTThrArGValThrTYTTrIeAspPrP 240

```

```

DB 667 GCGTGTGCCAGCCCAACCGCCGTGCATCTACACCCGGTGCACCTTACTTGGACTGG 726
QY 241 IleHISLeTYrValIProlyAspPro 249
DB 727 ATCCACCACTAATGTCCCAAAAAGCCG 753

```

```

RESULT 2
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
; GENERAL INFORMATION:
; APPLICANT: Mallit, Mark A.
; APPLICANT: Niles, Andrew L.
; APPLICANT: Haak-Frendscho, Mary
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human
; TITLE OF INVENTION: Beta-Trypsin and Method of Making Same
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Intellectual Property Department
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WISCONSIN
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,970A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leone, Joseph T.
; REGISTRATION NUMBER: 37,170
; REFERENCE/DOCKET NUMBER: 34506,073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..735

```

US-09-079-970A-1

```

Alignment Scores:
Pred. No.: 4.49e-148          Length: 735
Score: 1368.00              Matches: 244
Percent Similarity: 99.59%   Conservative: 0
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 98.21%         Indels: 0
DB: 3                         Gaps: 0

```

US-09-598-982c-21 (1-249) x US-09-079-970A-1 (1-735)

```

QY 5 IleValGIYGIInGUaLaPProArGserIyTPrPProTgInValSerLeuArGVal 24
DB 1 ATCGTCGGGGGTCAAGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGTGGAGGCTGAGAGTTC 60
QY 25 HISGLYProTYTTPMethHISpHeCySGIYserLeuIleHISProGlnITPValIleu 44
DB 61 CACGCCCACTAATGACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGCTG 120

```

QY 45 ThrAlaAlaAlaCysValGlyProAspValIlysaApleuAlaAlaLeuArgValGlnIleu 64
 DB 121 ACCGAGGACACTGGTGGGACCGGACCGTCAAGGATGGGGCCCTCAGGGTGAACCTG 180
 QY 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuProValSerArgIleIleValHis 84
 DB 181 CGGGAGCAGCACTTACTACAGAGACCAAGCTGCTGGCCGGTACAGAGATCATCTGTGCAC 240
 QY 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluPro 104
 DB 241 CCACAGTTCCTACACCGCCAGATCCGAGGCGCATTCGCCCTCTGAGAGTGGAGGACCGG 300
 QY 105 ValIlyValSerSerHisValHisThrValThrLeuProProAlaSerGlyThrPhePro 124
 DB 301 GTGAAGGCTCTCCAGCCAGCTCCACCGGTCACTCCCTCCCTCCAGAGCCCTTCCCC 360
 QY 125 ProGlyMetProCysTrpValThrGlyTyrGlyAspValAspAsnAspGluArgLeuPro 144
 DB 361 CCGGGGATGCCCTGCTGGGTCACTGGCTGGGGCGATGTGGACAAATGATGAGGCGCTCCCA 420
 QY 145 PropProPheProLeuIlyValGlnValIlyValProIleMetGluAsnHisIleCysAlaPala 164
 DB 421 CCGGCATTCTCTGAGAGCGAGGTGAAGGTCCCCATTAATGGAAAACCACTTTGTGACGCA 480
 QY 165 LysTyrHisIleuGlyValIleTyrThrGlyIlyAspAspValAlaGlyIleValArgAspAspMetLeu 184
 DB 481 AAATGCCACCTTGGGGCTTACACGGGAGACGACGTCCTCCGATGTCTCCGACCAATGTCTG 540
 QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyIlyProLeuValCys 204
 DB 541 TGTGCGGGGAGACACCGGAGGAGCTCATTCACGGGCGAGCTCCGGAGGGCCCTTGGTGTCC 600
 QY 205 LysValIleuGlyTyrThrIleuGlnIleGlyValIleValIleSerTrpGlyGluGlyCysAlaGln 224
 DB 601 AAGTGAATGGACCTGGCTGACGGGGCGGGCGTGTCAAGCTGGGGCGAGGGCGTGTGCCCG 660
 QY 225 ProAsnArgProGlyIleTyrThrArgValIleTyrTyrLeuAspTrpIleHisIleTyr 244
 DB 661 CCCAACCGGCTGGGATCTACACCGCTGTCACTTACTTGGACTGGATCCACCACTAAT 720
 QY 245 ValProIlyValPro 249
 DB 721 GTCCCCAAAAGCCCG 735
 RESULT 3
 US-09-016-366A-20
 ; Sequence 20, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Huang, Chitu
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Molt, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,366A
 ; FILING DATE: January 30, 1998
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/037,090

FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELE: _____
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1128 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-016-366A-20
 Alignment Scores:
 Pred. No.: 8,69e-148 Length: 1128
 Score: 1368.00 Matches: 244
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 98.21% Indels: 0
 DB: 2 Gaps: 0
 US-09-598-982c-21 (1-249) x US-09-016-366A-20 (1-1128)
 QY 5 ILeValGlyGlnGlnIleValIleProArgSerIlyTrpProTrpGlnValSerLeuArgVal 24
 DB 89 ATCGTGGGGTCAAGAGAGCCCGCCAGAGCAAGTGGCCCTGAGAGTGAAGCTGCA 148
 QY 25 HisGlyProIlyTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValIleu 44
 DB 149 CAGCGCCCATCTGGATGTCACCTTCTGCGGGGCTCCCTCATCCACCCCAAGTGGGTGTG 208
 QY 45 ThrAlaAlaAlaCysValGlyProAspValIlysaApleuAlaAlaLeuArgValGlnIleu 64
 DB 209 ACCGAGGCGCACTGGTGGGACCGGACCTCAAGATCTGGCCCGCTCAGGGTGAACCTG 268
 QY 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuProValSerArgIleIleValHis 84
 DB 269 CGGGAGCAGCACTTACTACAGAGACCAAGCTGCTCCGACGAGATCATCTGTGCAC 328
 QY 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluPro 104
 DB 329 CCACAGTTCCTACACCGCCAGATCCGAGGCGCATTCGCCCTGGAGAGGAGGACCGG 388
 QY 105 ValIlyValSerSerHisValHisThrValThrLeuProProAlaSerGlyThrPhePro 124
 DB 389 GTGAAGTTCCTCCAGCCAGCTCCACAGGTCACCTTCCCTCCCTCAGAGACTTCCCTCC 448
 QY 125 ProGlyMetProCysTrpValThrGlyTyrGlyAspValAspAsnAspGluArgLeuPro 144
 DB 449 CCGGGGATGCCCTGCTGGGTCACTGGCTGGGGCGGATGTGGACAAATGATGAGGCGCTCCCA 508
 QY 145 PropProPheProLeuIlyValGlnValIlyValProIleMetGluAsnHisIleCysAlaPala 164
 DB 509 CCGGCATTCTCTGAGAGCGAGGTGAAGTCCCCATTAATGGAAAACCACTTTGTGACGCA 568
 QY 165 LysTyrHisIleuGlyValIleTyrThrGlyIlyAspAspValAlaGlyIleValArgAspAspMetLeu 184
 DB 569 AAATGCCACCTTGGGGCTTACACGGGAGACGACGTCCTCCGATGTCTCCGAGACACTGCTG 628
 QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyIlyProLeuValCys 204
 DB 629 TGTGCGGGGAGACACCGGAGGAGCTCATGCGCAGGCGAGCTCCGGAGGGCCCTTGGTGTGC 688
 QY 205 LysValIleuGlyTyrThrIleuGlnIleGlyValIleValIleSerTrpGlyGluGlyCysAlaGln 224
 DB 689 AAGGTGAATGGACACCTGGCTCAGCGGGCGTGTGTCACTGGGCGAGAGGCGTGTGCCCG 748
 QY 225 ProAsnArgProGlyIleTyrThrArgValIleTyrTyrLeuAspTrpIleHisIleTyr 244

Db 749 CCCAACCGGCTGGAGCTACACCCGCTGCACCTACTACTGAGATGCCACCACTAT 808
 Qy 245 ValProlylsylPro 249
 Db 809 GTCCCAAAAAGCCG 823

RESULT 4
 US-08-978-404B-15
 ; Sequence 15, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 TITLE OF INVENTION: FIBRINOGEN
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESS: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:

INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1128 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-978-404B-15
 Alignment Scores:
 Pred. No.: 8.69e-148 Length: 1128
 Score: 1368.00 Matches: 244
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 98.21% Gaps: 0
 DB: 2

US-09-598-982c-21 (1-249) x US-08-978-404B-15 (1-1128)
 Qy 5 lIeValGIyGIyGInGIuAlaPrOArGSeLySTrPrOTrPGInValSerIeUArYVal 24
 Db 89 ATCGTGGGGGTCCAGGAGGCCCCCGAGCAAGGTGGCCCTGGCAGGTGAGCCCTGAGAGTC 148
 Qy 25 HIsGlyPrOTrYrTrImeHIsPheCySgIyGlySerIeUleHIsPrOgInTrYValIeu 44
 Db 149 CACGGCCCACTACTGAGTGGACTTCTGGCGGGGCTCCCTATPCACCCCAAGTGGGGTCTG 208
 Qy 45 ThrAlaAlaAlaCyEValGIyPrOAsPValIlyAsPleUAlAlaLeUArYValGIu 64
 Db 209 ACCGGAGCCCACTGGGTGGGAGCCGAGCGTCAAGAGATCTGGCCGCTCAAGGGTGAACACTG 268

Qy 65 ArgGIuGIInHIsLeUyTrYrGIInAsPGIInLeUeUProValSerArYIleIleValHIs 84
 Db 269 CGGGAGCGACACCTACTACTACAGGACCAAGTGGCTGGCCGGTCCAGGAGATCACTCGTGCAC 328
 Qy 85 PrOgInPheYrTrThAlaGIInIleGIyAlaAaPrIleAlaLeUeUInLeUeUInGIu 104
 Db 329 CCACAGTCTACACCGCCCAAGTGGAGCGGACATGGCTGTGGAGCTGGAGGAGCGC 388
 Qy 105 ValIySValSerSerHIsValHIsThrValThrIeUProProAlaSerGIuThrPhePr 124
 Db 389 GTMAAGTCTCCAGGCAAGTCCACACGCTCACCCCTGCCCTCCAGAGACCTTCCCC 448
 Qy 125 PrOgIyMeTProCySTrPrValThrGIyTrPGIyAsPValAsPAsPPrGIuArGIeUPro 144
 Db 449 CCGGGGATGCGGTGGTGGTCACTGGCTGGGGCGATGTGGGCAATGATGATGAGCGCCCTCCA 508
 Qy 145 PrOProPheProIeUySValIyValIyValIProIleMeTGIuAeNHIsIleCySAsPAla 164
 Db 509 CCGCCATTTCTCTGAGACAGGTGAGAGTCCCATATGAGAAACCAATTTGTGACCGCA 568
 Qy 165 LysTrYrHIsLeUyAlaIyTrThrGIyAsPAsPValArGIleValArGAsPAsPMeTleU 184
 Db 569 AAATACCACTTTGGCGCTTACACGGGAGACGACCGTCCGATCTCTCCAGACGACATGCTG 628
 Qy 185 CySAlaGIyAsPTrHArGArGAsPSeTcySgInGIyAsPSeTGIyGIyPrOleUValCyS 204
 Db 629 TGTGGCGGAAACACCCCGAGGACCTCATGCCAGGGCGACTCCGGGAGGGCCCTGGTGTGC 688
 Qy 205 LysValAsnGIyThrTrPLeUeUAlaGIyValIyValIserTrPGIyGIuGIyCySAlaGI 224
 Db 689 AAGGTGAATGGACCTGGCTGACGGCGGGCGTGTCACTGGGGGAGGGCTGTGGCCAG 748
 Qy 225 PrOAsnArGPrOgIyIleYrThrArYValThrYrYrIeUAsPTrPIleHIsIeYr 244
 Db 749 CCCAACCGGCTGGATGATACCCCGGTGTCACTACTACTGATCTGGATCCACCACTAT 808
 Qy 245 ValProlylsylPro 249
 Db 809 GTCCCAAAAAGCCG 823

RESULT 5
 US-09-016-366A-18
 ; Sequence 18, Application US/09016366A
 ; Patent No. 5955451
 ; GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 TITLE OF INVENTION: INHIBITORS
 NUMBER OF SEQUENCES: 65
 CORRESPONDENCE ADDRESS:
 ADDRESS: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,366A
 FILING DATE: January 30, 1998
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,090
 FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULAR TYPE: cDNA
 US-09-016-366A-18

Alignment Scores:
 Pred. No.: 3,32e-147 Length: 1137
 Score: 1363.00 Matches: 243
 Percent Similarity: 99.18% Conservative: 0
 Best Local Similarity: 99.18% Mismatches: 2
 Query Match: 97.85% Indels: 0
 DB: Gaps: 0

US-09-598-982C-21 (1-249) x US-09-016-366A-18 (1-1137)

5 lIeValGlyGlnGlnGlnAlaProArgSerLySTPProTTPGlnValSerLeuArgVal 24
 87 ATCGTCGGGGGTCAAGAGGCCCCCGAGCAAGTGGCCCTGGCAGGTGAGAGTGC 146
 25 HIsGlyProTyrTTPMethHisPheCysGlyGlySerLeuIleHisProGlnTTPValLeu 44
 147 CACGGCCCCCTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCGAGTGGTGTG 206
 45 ThrAlaAlaAlaCysValAlaGlyProArgValLysAspLeuAlaAlaLeuArgValGlnLeu 64
 207 ACCGAGCGCACTGGGTGGAGCCGACGTCAGAGATCTGGCCGCTCAGGGGTGCACTG 266
 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
 267 CCGGAGGAGCACTCTACTACAGAGCAAGCCTGCGCGGTGAGCAGATCATCGTGCAC 326
 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlnLeuGlnPro 104
 327 CCACAGTCTTACACCGCCAGATCCGAGGAGGATCCGCTCTGAGAGTGGAGGAGCCG 386
 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnIleThrPhePro 124
 387 GTAAACGTCCTCAAGCCAGCCAGCGGTCACCCCTGCCCCCTCAGAGACCTTCCCC 446
 125 ProGlyMetProCysTTPValThrGlyTTPGlyAspValAspAsnAspGlnUArgLeuPro 144
 447 CCGGGGATGCGCTGCTGGGTCACTGGCTGGGGCGATGTCGACAAATGATGAGGCGCTCCA 506
 145 ProProPheProLeuLysGlnValLysValProIleMetGlnAsnHisIleCysAspAla 164
 507 CCGCCATTCTCTGTAAGCAGGTTGAAAGTCCCATTAATGAAACCACTTTGTGACCGCA 566
 165 LysTyrHisIleLeuGlyAlaTyrThrGlyLysAspValArgIleValArgAspAsnMetLeu 184
 567 AAATACCACTTGGGCTTACAGGAGAGCGAGCTCCGATGCTCCGATGACCACTGCTG 626
 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 627 TCTGCCGGGAAACACCCCGAGGAGCTCATCCCGAGGCGACTCCGGAGGGGCCCTGGTGTCC 686
 205 LysValAsnGlyThrTTPLeuGlnAlaGlyValValSerTTPGlyGlnGlyCysAlaGln 224
 687 AAGGTGAATGGCACTGGCTGCAAGGGGGCGTGTACAGCTGGGGGGGAGGGCGTGTCCCG 746
 225 ProAsnArgProGlyIleTyrThrPArgValThrTyrTyrLeuAspTTPIleHisIleTyr 244
 747 CCAACCGGCTGGCATCTTACACCGGTGTCACTTACTTACTTGAATCCACCACTAAT 806
 245 ValProLysLysPro 249

DB 807 GTCCCAAAAAGCCG 821

RESULT 6

US-08-978-404B-13

Sequence 13, Application US/08978404B

Patent No. 5968782

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.

TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES

TITLE OF INVENTION: FIBRINOGEN

NUMBER OF SEQUENCES: 74

CORRESPONDENCE ADDRESS:

ADDRESS: Wolf, Greenfield & Sacks, P. C.

STREET: 600 Atlantic Avenue

CITY: Boston

STAT: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978, 404B

FILING DATE: 25-NOV-97

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,354

FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 36,637

REFERENCE/DOCKET NUMBER: B0801/7090

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

TEXT:

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 1137 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-978-404B-13

Alignment Scores:

Pred. No.: 3,32e-147 Length: 1137

Score: 1363.00 Matches: 243

Percent Similarity: 99.18% Conservative: 0

Best Local Similarity: 99.18% Mismatches: 2

Query Match: 97.85% Indels: 0

DB: Gaps: 0

US-09-598-982C-21 (1-249) x US-08-978-404B-13 (1-1137)

5 lIeValGlyGlnGlnGlnAlaProArgSerLySTPProTTPGlnValSerLeuArgVal 24
 87 ATCGTCGGGGGTCAAGAGGCCCCCGAGCAAGTGGCCCTGGCAGGTGAGAGTGC 146
 25 HIsGlyProTyrTTPMethHisPheCysGlyGlySerLeuIleHisProGlnTTPValLeu 44
 147 CACGGCCCCCTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCGAGTGGTGTG 206
 45 ThrAlaAlaAlaCysValAlaGlyProArgValLysAspLeuAlaAlaLeuArgValGlnLeu 64
 207 ACCGAGGAGCACTCGTGGGAGCCGAGCGTCAAGGATCTGGCCGCTCAGGGGTCAACTG 266
 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
 267 CCGGAGGAGCACTCTACTACAGAGCAAGCCTGCGCGGTGAGCAGATCATCGTGCAC 326
 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlnLeuGlnPro 104

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1081 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-016-366A-22

Alignment Scores:
Pred. No.: 4.76e-145 Length: 1081
Score: 1344.00 Matches: 241
Percent Similarity: 98.37% Mismatches: 4
Best Local Similarity: 98.37% Indels: 0
Query Match: 96.48% Gaps: 0
DB: 2

US-09-598-982c-21 (1-249) x US-09-016-366A-22 (1-1081)
QY 165 LysTyrHisLeuGlyValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 567 AAATACCACTCTGGCCCTCAGCGGAGACGACGTCCTCCGATCGTCCGTAGCAATGCTG 626
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
DB 627 TGTGCCGGGAACCCCGAGGAGCACTCATGCCAGGGCGACTCCGGAGGGCCCTGGTGTGC 686
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
DB 687 AAGGTAAATGGCCCTGGCTCAGCGGCGGTGGTCACTGGGGCCAGGGCTGTGCCAG 746
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIleTyr 244
DB 747 CCCAACGGGCTGGCATCTACACCCGCTGTCACTTACCTTGGACTGGATCCACCACTTAT 806
QY 245 ValProIlyIysPro 249
DB 807 GTCGCCAAAAAAGCCG 821

```

```

RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1081 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-016-366A-22

Alignment Scores:
Pred. No.: 4.76e-145 Length: 1081
Score: 1344.00 Matches: 241
Percent Similarity: 98.37% Mismatches: 4
Best Local Similarity: 98.37% Indels: 0
Query Match: 96.48% Gaps: 0
DB: 2

US-09-598-982c-21 (1-249) x US-09-016-366A-22 (1-1081)
QY 5 IleValGlyGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
DB 68 ATCGTTGGGGGTCAAGAGCCCGCCAGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTGC 127
QY 25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
DB 128 CGGACCGATCTGATGACTTCTGGGGGCTCTCATATCCACCCCACTGGGTGCTG 187
QY 45 ThrAlaAlaAlaCysValGlyProAspValIlyAspLeuAlaAlaLeuArgValGlnLeu 64
DB 188 ACCGAGCGCACTGGCTGGAGACCGGACCGTCAAAGATCTGGCCCGCTCAGGGTGGCACTG 247
QY 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
DB 248 CGGGAGCAGCACTCTACTACCGAGCACGCTGCGCCGCTCAGGAGATCATCTGTGAC 307
QY 85 ProGlnPheTyrThrAlaGlnIleGlyValAspIleAlaLeuLeuGlnLeuGlnIlePro 104
DB 308 CCACAGTTCTACACCGCCCAATCCGAGCGGACATCGCCCTGCTGAGACTGGAGAGCCG 367
QY 105 ValIysValSerSerHisValHisIleThrValThrLeuProProAlaSerGlyThrPhePro 124
DB 368 GTGAAGGTCTCCAGCAGCTCCACACCGGTCAACCCCTGCCCTCAGAGACTCTTCCC 427
QY 125 ProGlyMetProCysTrpValThrGlyTyrGlyAspValAspAspAspGluArgLeuPro 144
DB 428 CCGGGGATGCGCTGTGGTCACTGCTGGGGCGATGTGGACAAATGATGAGGGCCCTCCA 487
QY 145 ProProPheProLeuLeuGlnValIlyValProIleMetGluAsnHisIleCysAspAla 164
DB 488 CCGCAATTTCTCTGAAAGAGGTAAAGTCCCAATATGAAACCAACATTTGTGACGCA 547
QY 165 LysTyrHisLeuGlyValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 548 AAATACCACTTGGCGCTTACACCGGAGACGACGTCGCCATCTGTCGTGACGACATGCTG 607
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
DB 608 TGTGCCGGGAACCCCGAGGAGCACTCATGCCAGGGCGACTCCGGAGGGCCCTGGTGTGC 667
QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
DB 668 AAGGTGAATGGCACTTGGCTCAGGGCGGGCGGTGTCAGCTGGGGGAGGGGCTGTGCCAG 727
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIleTyr 244
DB 728 CCCAACGGGCTGGATCTACACCCGCTGTCACTTACTTGGACTGGATCCACCACTAT 787
QY 245 ValProIlyIysPro 249
DB 788 GTCGCCAAAAAAGCCG 802

```

```

RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B

```



```

QY 65 ArgGluGlnHisLeuTyrTrpGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
DB 248 CCGGAGCAGACCTCTACTACCAAGGACCAAGCTCTGCTCCGGTCCAGAGATCATCGTGACC 307
QY 85 ProGlnPhePheTrpThrAlaGlnIleGlyValAspIleAlaLeuLeuGluGluGluPro 104
DB 308 CCAAGGTTCTACACCCGCGCATCCGAGATCCGAGATCGCCCTGTGGAGCTGGAGGCGCCG 367
QY 105 ValIleValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
DB 368 GTGAAGGTCTCCACCAAGCCACAGGTCACCCCTGCTCCCAAGACCTTCCCTCC 427
QY 125 ProGlyMetPheProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluValGlnLeuPro 144
DB 428 CCGGGGATGCGGTGTGGGTCATCTGGCTGGGGAGATGTGGACAATGATGAGCGCTCCCA 487
QY 145 ProProPheProLeuLeuGlnValIleValysValProIleMetGluAsnHisIleCysAspAla 164
DB 488 CCGCCATTTCTCTGAAAGCAAGTAAAGTCCCAATATGAAACCAACATTGTGTAGCGCA 547
QY 165 IysTyrHisLeuGlyValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 548 AAATACACACCTTGGCGCCTACAGGAGGAGAGAGAGATCGCCGATTCGTCGATGACATGCTG 607
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
DB 608 TGTGCCGGGAACACCCGGAGGAGACTCATGCCAGGGGCACTCCGGAGGGCCCTGGTGTGTC 667
QY 205 IysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
DB 668 AAGGTAAATGCGACCTGGCTGACAGGGCGGGGTGTGCACTGGGGCCAGAGGCTGTGGCCAG 727
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTrpLeuAspTrpIleHisIstYr 244
DB 728 CCCAACCGGCTGGCATCTACACCCGCTGTCTACTACTAGCTTGGACTGGATCCACCTAT 787
QY 245 ValProIysIysPro 249
DB 788 GTCGCCAAAAGGCCG 802

```

```

RESULT 10
US-09-016-366A-16
Sequence 16, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSER: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093

```

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-720-3500
? TELEFAX: 617-720-2441
?
?
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1154 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
US-09-016-366A-16
Alignment Scores:
Pred. No.: 6.2e-137 Length: 1154
Score: 1274.00 Matches: 226
Percent Similarity: 94.69% Conservative: 6
Best Local Similarity: 92.24% Mismatches: 13
Query Match: 91.46% Indels: 0
DB: 2 Gaps: 0

```

```

US-09-598-982C-21 (1-249) x US-09-016-366A-16 (1-1154)
QY 5 IIEValGIYGLYGLInGluAlaProArgSerIysTrpProTrpGlnValSerLeuArgVal 24
DB 108 ATCGTCCGGGGGTCCAGAGAGGCCCCAGAGCAAGTGGCCCTGGCAGGTGAGCTTGAGAGTC 167
QY 25 HIsGlyProTyrTrpMetHisPheCysGlyIysSerLeuIleHisProGlnTrpValIleu 44
DB 168 CCGGACCCGATCTGGATGCACTTCTGCGGGGGCTCCCTCAITCCACCCCAAGTGGTGTCTG 227
QY 45 ThrAlaAlaAlaCysValGlyProAspValIysAspLeuAlaAlaLeuArgValGlnLeu 64
DB 228 ACCCGGGGCACTGGCTGGAGCCGGAGCCGATGCAAGATCTGGCCCACTCCAGGGGTGAAGCTG 287
QY 65 ArgGluGlnHisLeuTyrTrpGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
DB 288 CCGGAGCAGCACCTGTAACAACAGAGCAAGCTGGCCAGTCAAGAGATCATCTGTGAC 347
QY 85 ProGlnPheTyrThrAlaGlnIleGlyValAspIleAlaLeuLeuGluGluGluPro 104
DB 348 CACAGGTTCTACATGATCACTGACAGTGGAGTATTCGCCCTCTGGACTGGAGAGAGCC 407
QY 105 ValIysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
DB 408 GTGAACATCTCCAGCCGGCTCCACAGGTCATGTGCTGCCCTCCGGAGACCTTCCCC 467
QY 125 ProGlyMetPheProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
DB 468 CCGGGATGCGGTGTGGCTGCTGCTGCTGGCGGATGTGGACAATGATGAGCCCTCCCA 527
QY 145 ProProPheProLeuLeuGlnValIleValysValProIleMetGluAsnHisIleCysAspAla 164
DB 528 CCGCCATTTCCCTGTAAGCGGTGAAGTCCCAATATGAAACCAACATTTGTGACGA 587
QY 165 IysTyrHisLeuGlyValAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
DB 588 AAATACCACTTGGCGCCTTACAGGGAAGCAGCCTCCGATCATCCGTGAGCAATGCTGC 647
QY 185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
DB 648 TGTGCCGGGAACACCCGGAGGAGCTCTGCAAGGGCGACTGTGGAGGGCCCTGTGTGTC 707
QY 205 IysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
DB 708 AAGGTGAATGCACTGGCTACAGCGGGCGGTGTCACTGGAGAGAGAGGCTGTGCCAG 767
QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTrpLeuAspTrpIleHisIstYr 244
DB 768 CCAACCGGCGTGGATCAACCCGCTGTCACTACTTGTGAGATGATCCACCACTAT 827
QY 245 ValProIysIysPro 249

```


DB 828 GTCCCAAAAAGCGG 842

RESULT 11
US-08-978-404B-11
Sequence 11, Application US/08978404B
Patent No. 5968782

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
NUMBER OF INVENTIONS: FIBRINOGEN 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
ADDRESS: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996

REGISTRATION NUMBER: 36,637
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441

TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1154 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-978-404B-11
Alignment Scores:
Pred. No.: 6.2e-137 Length: 1154
Score: 1274.00 Matches: 226
Percent Similarity: 94.69% Conservat: 6
Best Local Similarity: 92.24% Mismatches: 13
Query Match: 91.46% Indels: 0
DB: 2 Gaps: 0

US-09-598-982C-21 (1-249) x US-08-978-404B-11 (1-1154)

QY 5 l l e v a l g l y g l i n g l u a l a p r o a r g s e r l y t t p r o t r p g i n v a l s e r t e u a r g v a l 24

DB 108 a t t c g t c g g g g g t c a g a g a c c c c c c c a g a c c a a g t g g c c c t g c a g a t g a g a g t t c 167

QY 25 h i s g l y p r o t y r t r p m e t h i s p h e c y s g i y g l y s e r t e u l l h i s p r o d i n t r p v a l l e u 44

DB 168 c g c g a c c a t c t g a t t g a c t t t g c g g g g g t c c c t a t a c a c c c a g t g g g t g c t g 227

QY 45 t h r a l a a l a a l a c y e v a l g l y p r o a r p v a l l y s a s p l e u a l a l a e u a r g v a l g i n e u 64

DB 228 a c c g g g g c a c t g c c t g g a c c g a c g r c a a g a n t c t g c c a c c t c a g g g t g c a a c t g 287

QY 65 a r g g l u g l i n h i s l e u t y r t y r g i n a p p g i n l e u p r o v a l s e r t a r g i l l e i l e v a l h s 84

DB 288 c g g a g a c a c a c c t t a c t a c c a g a c c a c g t g c t g c c a g a c a g a t g a t t c c t g c g c a c 347

QY 85 p r o g i n p h e t y r t h r a l a g i n l e g i l y a l a a p p l e a l a l e u e u g l u e u g i u g i u p r o 104

DB 348 c c a c a g t t t r a c a t c a t c a c a g a c t g a g c g g a t t c c c t g c t g a g a g c t g a g a g a g c c 407

QY 105 v a l l y s v a l s e r s e r t h i s v a l h i s t h r v a l t h r l e u p r o p o a l a s e r g i u t h r p h e p r o 124

DB 408 g t g a a c a t t c t c a c c c g g t c a c a c c g t c a t g c t g c c c c t g c c t g c g a n a c c t t c c c c 467

QY 125 p r o g i y m e t p r o c y e t t r p v a l t h r g l y t r g l y a s p v a l a s p a n a a p p g l u a r g l e u p r o 144

DB 468 c g g a g a a g c c g t g t c t g g t a c t g g c t g g g g c a t g g a c a n t a g n i a g a c c c c t c c c a 527

QY 145 p r o p r o p h e p r o l e u y g l n v a l l y v a l p r o i l e m e t g l u a e m h i s i l e c y a s a p a l a 164

DB 528 c c c c a t t t c c c g a a g a c g a g t a g a g t c c c a n t a n t g a a a a a c a c a t t t g g a g a c g c a 587

QY 165 l y s t y r t h i s l e u g l y a l a y r t h r g l y a s p a s p v a l a r g i l e v a l a g a s p a s p a m e l e u 184

DB 588 a a a t r a c c a c t t g g c g c t t a c a c c g g a g a c a c a c t c c g c a t c a t c c g t g a c g a t a t g c t g 647

QY 185 c y a l a g i y a e n t h r a r g a r g a s p s e r c y e s g i n g l y a s p s e r g i y g l y p r o l e u v a l c y s 204

DB 648 t g t g c c g g a a c a c g c a g a g g a c t c c t g c a a g g c g h c t t g g a g g c c c t g g t g t g c 707

QY 205 l y s v a l s e n g i y t h r t r p l e u g l n a g l y v a l s e r t p g l y g l u g l y c y a l a g i n 224

DB 708 a a g t g a n a t g c a c a c t g c t r a c a g c g g c c t g t c a g c t g g a n g a g g c t g t g c c c a g 767

QY 225 p r o a n a r g p r o g i y i l e t y r t h r a r g v a l t h r t y r l e u a a p t r p l i h i s i s t y r 244

DB 768 c c c a a c c g c c t g c a r t a c a c c g t g t c a c t a c t t g a c t t g a c t g a t t c c a c c a c t a t 827

QY 245 v a l p r o l y s l y s p r o 249

DB 828 g t c c c a a a a a g c g g 842

RESULT 12
US-09-016-366A-14

Sequence 14, Application US/09016366A
Patent No. 5955431

GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
ADDRESS: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997

ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:

INFORMATION FOR SEQ ID NO: 14:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-016-366A-14

```

```

Alignment Scores:
Pred. No.: 7 31e-116 Length: 1108
Score: 1091.00 Matches: 190
Percent Similarity: 86.01% Conservative: 19
Best Local Similarity: 78.19% Mismatches: 34
Query Match: 78.32% Indels: 0
Gaps: 0

```

US-09-598-982C-21 (1-249) x US-09-016-366A-14 (1-1108)

```

OY 5 ILeValGlyGlnGlnuAlaProArGSerLysTrpProTrpGlnValSerLeuArGVal 24
Db 127 ATCGTGGAGGACATGATGAGGCTTCTGAGAGTAAGTGGCCCTGGGCGAGGCTGAGATTT 186
OY 25 HisGlyProTyrrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValIleu 44
Db 187 AAATTAACCTACTGAGTACATTTCTCGGAGGCTCTCTCATCCACCCAGTGGGCTC 246
OY 45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArGValGlnLeu 64
Db 247 ACTGCGGCACACCTGTGTGGACCGGCACATCAAAAGCCCAAGCTCTCCGGGTGACGCTT 306
OY 65 ArgGlnGlnHisLeuTyrrTyrrGlnAspGlnLeuProValSerArGLeIleValHis 84
Db 307 CGTGACACAGATATTAATCTATGAGGACCAAGCTCTCTTTGAACCGGATCGTGTGCAC 366
OY 85 ProGlnPheTyrrThraGlnIleGlyAlaAspIleAlaLeuLeuGlnLeuGlnIleuPro 104
Db 367 CCCCACTATTACAGGGCGGGTGGGACAGCTGCTGCTGGAGCTTGGAGTCCCT 426
OY 105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
Db 427 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCTGCTGGAGCTTGGAGTCCCT 486
OY 125 ProGlyMetProCysTrpValThrGlyTyrrGlyAspValAspAsnAspGluArGLeuPro 144
Db 487 CCTGGGACATCGTGTGGTGGACAGGCTGGGCGGACATTAATGATGACGAGCTTCCCA 546
OY 145 ProProPheProLeuLysGlnValLysValProIleMetGlnLysHisIleCysAspAla 164
Db 547 CTTCTTATCTGTGAAGCAAGTGAAGTCCCATTTGTGAAAACAGCCCTGTGTGACCGG 606
OY 165 LysTyrrHisLeuGlyAlaTyrrThrGlyAspAspValArGLeIleValArGAspAspMetLeu 184
Db 607 AAGTATCACACCTGCGCTCTACACGGGAGATGATTTTCCCATTTCCATGATGGCATGCTG 666
OY 185 CysAlaGlyAsnThrArGArGAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
Db 667 TGTGCTGGAATATCCAGGAGAGACTCTCTCCAGGGGATTCAGGGGGCCACAGTGTCTGG 726
OY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 224
Db 727 AAGGTAGAGGATACCTGGCTGTGAGGACGAGAGTGTGACTGCTGGGATGAGGGCTCCGACAG 786
OY 225 ProAsnArGProGlyIleTyrrThrArGValThrTyrrTyrrLeuAspTrpIleHisHisTYR 244
Db 787 CCGAACAAGCCTGGACATCTACACCCGGGTGACATTACTTAAGTACGAGATCCACCGCTAT 846
OY 245 ValProLys 247
Db 847 GTCCTGTAG 855

```

```

; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/70930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-598-982C-21 (1-249) x US-08-978-404B-20 (1-1108)
;
Alignment Scores:
Pred. No.: 7 31e-116 Length: 1108
Score: 1091.00 Matches: 190
Percent Similarity: 86.01% Conservative: 19
Best Local Similarity: 78.19% Mismatches: 34
Query Match: 78.32% Indels: 0
Gaps: 0

```

RESULT 13
US-08-978-404B-20
; Sequence 20, Application US/08978404B

427 GTGAATGCTCCACCATATATCCACCACCCAMTCCCTCCCTCCCTCCGAGACCTTCCCC 486
 QY 125 ProglYmeCProCYSTRVAlThrGlyTRpGlyVAspVAlAspAspGluAArgLeuPro 144
 DB 487 CCTGGACATCGTGTGGGTGACAGGCTGGGGCGCATTTGATTAAGACGACCTTCCCA 546
 QY 145 PropPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
 DB 547 CCTCTTATCTCTGAAAGCAAGTAAAGTTCCCATTTGGGAAAACAGCCCTGTGTGACCG 606
 QY 165 LysTyrHisIleLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
 DB 607 AAGTACCAACATCGCTGCTACACGGGAGATGATTTTCCCATTTGCTCATGATGACATGCTG 666
 QY 185 CysAlaGlyAsnThrArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 DB 667 TGTGCTGGAAATACAGAGAGACTCTCCAGGGCGATCCAGGGGGGCCACTGGTCTGCG 726
 QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlyCysAlaGln 224
 DB 727 AAGTGAAGGCTTACTGCTGGCTGACAGGACAGAGTGTCAAGCTGGGGGTGAGGGCTCCGCA 786
 QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
 DB 787 CCAACAGCCTGGCATCTACACCCGGGTGACATATCTTATGACTGTGATCCACCGCTAT 846
 QY 245 ValPProLys 247
 DB 847 GTCCCTGAG 855

RESULT 14
 US-08-978-404B-7
 ; Sequence 7, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1219 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-978-404B-7

Alignment Scores:
 Pred. No.: 5,93e-114 Length: 1219
 Score: 1075.00 Matches: 188
 Percent Similarity: 84.77% Conservative: 18
 Best Local Similarity: 77.37% Mismatches: 37
 Query Match: 77.17% Indels: 0
 DB: 2 Gaps: 0

US-09-598-982c-21 (1-249) x US-08-978-404B-7 (1-1219)

QY 5 ILeValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
 DB 273 ATGTTGGGGGACAGAGGGGACTGGGAAACAAGGCCCTGGCAGGTGAGGCTTGGTCC 332
 QY 25 HisGlyProTyrTrpMetHisArgPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
 DB 333 AATGAAACCTACTGAGAGCAATTTCTGGGGGCTCCCTCATCCACCCAGGTGGTCTC 392
 QY 45 ThrAlaIaAlaCysValGlyProAspValLysAspLeuAlaIleLeuArgValGlnLeu 64
 DB 393 ACCGGGACACTGTGTGGGACCGACTATTTGCTGATCCCAACAAGTTCAGAGTACAGCTT 452
 QY 65 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
 DB 453 CGAAAGGACTACTCTTATTAACAGACCACTGCTGGGTGTGAGCCGGATCATCACAC 512
 QY 85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlyLeuGlyGluPro 104
 DB 513 CCGACATTTCTATGTCACCCAGCAATATGGGGCGGACATCCCTCACTTGTGAGCTCAAGAACCTT 572
 QY 105 ValLysValSerSerHisIleValHisIleThrValThrLeuProProAlaSerGlyThrPhePro 124
 DB 573 GTAAACATTTCCAGCCAGTGTCCACCCCGTCTCCCTGCTTCCCTCAGAGACTTCCCC 632
 QY 125 ProGlymeCProCYSTRVAlThrGlyTRpGlyVAspVAlAspAspGluAArgLeuPro 144
 DB 633 TCAGGAACATTTGCTGGGTGACAGGCTGGGAAACAATCGACATGATGAGCCTGGCA 692
 QY 145 PropPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
 DB 693 CCGCATTTCCCTTGAAGAGGTGCAAGTTCCCTCGTGGAAAACAGCCTTGTGACTCG 752
 QY 165 LysTyrHisIleLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
 DB 753 AAGTATCAACAAGGTGTCTACACAGGGGACAACATCCCATTTGTCCGAGACGATGCTG 812
 QY 185 CysAlaGlyAsnThrArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
 DB 813 TGTGCTGGGAACGAAAGACAGACTCTCCGCAAGGTTGACTCCGAGAGACCTCGTGGTCTGC 872
 QY 205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlyGlyCysAlaGln 224
 DB 873 AAGGTAAACGGTACTGCTGCTGACAGGCAAGTGTGTGCTGGGGGTGAGGGCTGTGCTCG 932
 QY 225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
 DB 933 CCAACAGCCTGGCATCTACACTGGGTGACCTTATTAAGTGTGAGATCCACCGCTAT 992
 QY 245 ValPProLys 247
 DB 993 GTCCCAAG 1001

RESULT 15
 US-08-978-404B-1
 ; Sequence 1, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:

ADDRESSER: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-978-404B-1

Alignment Scores:

	Pred. No.:	1,21e-111	Length:	1031
Score:		1054.00	Matches:	186
Percent Similarity:		83.54%	Conservative:	17
Best Local Similarity:		76.54%	Mismatches:	40
Query Match:		75.66%	Indels:	0
DB:		2	Gaps:	0

US-09-598-982C-21 (1-249) x US-08-978-404B-1 (1-1031)

Db	531	CCACCATTTCTTTGGAAGGAGGTGCCAAGTCCCATTTAATAGAAAACCCCTTTGTGACTTG	590
Qy	165	LVSTYRHLSLEUQIYVAATYRTRGIVASPAAPVALARQILEVALARGAPAPMETLEU	184
Db	591	AAGTATACAAAGGTCTCATTCAGAGTGAACAATGTCCACATTGTCCGAGATGACATGCTG	650
Qy	185	CYSAIAGIYASNTHRARGHARGAPSERCYSGINGIYAESPSEYGLYPROLEUVVALCY	204
Db	651	TTGTGCTGGGAATGAAAGACATGACTCTGCAGGGGCACTCCGGAGGAACTCTGGGTCTGC	710
Qy	205	IYVVALAENGLYTHRTRLENGIHALAGIYVALVALSERTRPGIYGLUGIYCYBALAGI	224
Db	711	AAGGTAAAGACACTGTGCGGAGGACAGCGTGTCAAGCTGGGGAGGAGGTGTGCAAG	770
Qy	225	PROAENARGPROGLIYLETYRTHRARGVALTHRTRYRTRYLEVAAPTTPILHISHISTYR	244
Db	771	CCCAACAGGCCCTGGAGTCTTACACCCGGGTTCACTTATCTTGATGGATCCACCATAT	830
Qy	245	VALPROLYS 247	
Db	831	GTCCCAAG 839	

Search completed: August 27, 2005, 21:26:04
Job time : 183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:01:40 ; Search time 113 Seconds
(without alignments)
865.829 Million cell updates/sec

Title: US-09-598-982C-21
Perfect score: 1393
Sequence: 1 LDKRIVGGQEARSRKMPQV.....ITRVYTYLDWTHHVPKRP 249

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	/cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpa/US11A_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/2/pubpa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/2/pubpa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/2/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1368	98.2	254	14 US-10-131-409-74	Sequence 74, Appl
2	1368	98.2	254	14 US-10-139-854-74	Sequence 74, Appl
3	1368	98.2	254	15 US-10-150-813-74	Sequence 74, Appl
4	1368	98.2	254	15 US-10-150-811-74	Sequence 74, Appl
5	1368	98.2	256	14 US-10-131-409-92	Sequence 92, Appl
6	1368	98.2	256	15 US-10-139-854-92	Sequence 92, Appl
7	1368	98.2	256	15 US-10-150-813-92	Sequence 92, Appl
8	1368	98.2	256	15 US-10-150-811-92	Sequence 92, Appl
9	1368	98.2	264	14 US-10-131-409-104	Sequence 104, App
10	1368	98.2	264	15 US-10-139-854-104	Sequence 104, App
11	1368	98.2	264	15 US-10-150-813-104	Sequence 104, App

12	1368	98.2	264	15	US-10-150-811-104	Sequence 104, App
13	1368	98.2	275	10	US-09-813-432-44	Sequence 44, Appl
14	1368	98.2	275	15	US-10-352-684A-48	Sequence 48, Appl
15	1368	98.2	275	15	US-10-174-364-44	Sequence 44, Appl
16	1368	98.2	275	15	US-10-246-583-44	Sequence 44, Appl
17	1368	98.2	275	15	US-10-287-226-94	Sequence 94, Appl
18	1368	98.2	275	16	US-10-689-832-44	Sequence 44, Appl
19	1363	97.8	275	15	US-10-287-226-92	Sequence 92, Appl
20	1363	97.8	275	16	US-10-723-860-2690	Sequence 2690, Ap
21	1363	97.8	275	18	US-10-756-149-5451	Sequence 5451, Ap
22	1347.5	96.7	272	14	US-10-117-323-40	Sequence 40, Appl
23	1329	95.4	691	15	US-10-275-505-13	Sequence 13, Appl
24	1323	95.0	275	15	US-10-287-226-90	Sequence 90, Appl
25	1317	94.5	266	14	US-10-131-409-84	Sequence 84, Appl
26	1317	94.5	266	15	US-10-139-854-84	Sequence 84, Appl
27	1317	94.5	266	15	US-10-150-813-84	Sequence 84, Appl
28	1317	94.5	266	15	US-10-150-811-84	Sequence 84, Appl
29	1312	94.2	279	15	US-10-287-226-88	Sequence 88, Appl
30	1274	91.5	245	11	US-09-789-210-69	Sequence 69, Appl
31	1274	91.5	275	14	US-10-117-323-36	Sequence 36, Appl
32	1274	91.5	275	15	US-10-352-684A-46	Sequence 46, Appl
33	1270	91.2	231	10	US-09-898-837A-52	Sequence 52, Appl
34	1104	79.3	199	14	US-10-131-409-98	Sequence 98, Appl
35	1104	79.3	199	15	US-10-139-854-98	Sequence 98, Appl
36	1104	79.3	199	15	US-10-150-813-98	Sequence 98, Appl
37	1104	79.3	199	15	US-10-150-811-98	Sequence 98, Appl
38	1091	78.3	252	10	US-09-813-432-42	Sequence 42, Appl
39	1091	78.3	252	15	US-10-174-364-42	Sequence 42, Appl
40	1091	78.3	252	15	US-10-246-583-42	Sequence 42, Appl
41	1091	78.3	252	16	US-10-689-832-42	Sequence 42, Appl
42	1091	78.3	276	14	US-10-117-323-35	Sequence 35, Appl
43	1077	77.3	276	14	US-10-311-955-3	Sequence 3, Appl
44	1064	76.4	247	14	US-10-131-409-86	Sequence 86, Appl
45	1064	76.4	247	15	US-10-139-854-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-10-131-409-74
; Sequence 74, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
; APPLICANT: Mal'ankar' et al.
; TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CIPICOM1
; CURRENT APPLICATION NUMBER: US/10/131,409
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/898,954
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; PRIOR FILING DATE: 2000-09-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74

; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-131-409-74

Query Match 98.2%; Score 1368; DB 14; Length 254;
 Best Local Similarity 99.6%; Pred. No. 4.2e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPKRSMQVSLRKHGPRYMMHFCGSSLIHPPQWVLTAAACVGPDPVKDLAALRYQL 64
 |||||
 Db 10 IVGGQEARPKRSMQVSLRKHGPRYMMHFCGSSLIHPPQWVLTAAACVGPDPVKDLAALRYQL 69
 QY 65 REQHLYYQDQLLPSVRSIIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASSETFP 124
 |||||
 Db 70 REQHLYYQDQLLPSVRSIIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASSETFP 129
 QY 125 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 184
 |||||
 Db 130 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 189
 QY 185 CAGNTRRDSGCGDGGPILVCKVNGTWTLAGVSWGEGCAQPNRPGIYTRVYYLDWIMHHY 244
 |||||
 Db 190 CAGNTRRDSGCGDGGPILVCKVNGTWTLAGVSWGEGCAQPNRPGIYTRVYYLDWIMHHY 249
 QY 245 VPKKP 249
 |||||
 Db 250 VPKKP 254

RESULT 2
 US-10-139-854-74
 ; Sequence 74, Application US//10139854
 ; Publication No. US20030202971A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder, Kumar
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-675CON2
 ; CURRENT FILING DATE: 2002-12-02
 ; PRIOR APPLICATION NUMBER: 09/783,429
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 74
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-139-854-74

Query Match 98.2%; Score 1368; DB 15; Length 254;
 Best Local Similarity 99.6%; Pred. No. 4.2e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPKRSMQVSLRKHGPRYMMHFCGSSLIHPPQWVLTAAACVGPDPVKDLAALRYQL 64
 |||||
 Db 10 IVGGQEARPKRSMQVSLRKHGPRYMMHFCGSSLIHPPQWVLTAAACVGPDPVKDLAALRYQL 69
 QY 65 REQHLYYQDQLLPSVRSIIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASSETFP 124
 |||||
 Db 70 REQHLYYQDQLLPSVRSIIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASSETFP 129
 QY 125 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 184
 |||||
 Db 130 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 189
 QY 185 CAGNTRRDSGCGDGGPILVCKVNGTWTLAGVSWGEGCAQPNRPGIYTRVYYLDWIMHHY 244
 |||||
 Db 190 CAGNTRRDSGCGDGGPILVCKVNGTWTLAGVSWGEGCAQPNRPGIYTRVYYLDWIMHHY 249
 QY 245 VPKKP 249
 |||||
 Db 250 VPKKP 254

RESULT 3
 US-10-150-813-74
 ; Sequence 74, Application US//10150813
 ; Publication No. US20030224367A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-675CIP1CON2
 ; CURRENT FILING DATE: 2002-12-03
 ; PRIOR APPLICATION NUMBER: US/10/150,813
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 09/898,954
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 74
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-150-813-74

Query Match 98.2%; Score 1368; DB 15; Length 254;
 Best Local Similarity 99.6%; Pred. No. 4.2e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 130 PGMPCWVTGMDVNDERLPPFPPLKQVYPIIMENHI CDAKYHLGAYTGDVRIYRDDML 189
Qy 185 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 244
Db 190 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 249
Qy 245 VPKKP 249
Db 250 VPKKP 254

```

RESULT 4

```

US-10-150-811-74
; Sequence 74, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
; APPLICANT: Malyankar et al.
; TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675C1P2CON1
; CURRENT FILING DATE: 2002-05-17
; PRIOR FILING DATE: 2001-10-03
; PRIOR FILING DATE: 09/970,607
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-150-811-74

```

```

Query Match 98.2%; Score 1368; DB 15; Length 254;
Best Local Similarity 99.6%; Pred. No. 4.2e-122;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 IYGGGEAERSKMPQVSLRKHVGRPYWMMHFCGSSLIHPQWVLTAAACVGPVDVRLAALRVQL 64
Db 10 IYGGGEAERSKMPQVSLRKHVGRPYWMMHFCGSSLIHPQWVLTAAACVGPVDVRLAALRVQL 69
Qy 65 RBOHLYYDDQLLPPRSRIIVHPQFYTAQIGADIALLELEBPVKVSSHVTITLPPASSTFP 124
Db 70 RBOHLYYDDQLLPPRSRIIVHPQFYTAQIGADIALLELEBPVKVSSHVTITLPPASSTFP 129
Qy 125 PGMPCWVTGMDVNDERLPPFPPLKQVYPIIMENHI CDAKYHLGAYTGDVRIYRDDML 184
Db 130 PGMPCWVTGMDVNDERLPPFPPLKQVYPIIMENHI CDAKYHLGAYTGDVRIYRDDML 189
Qy 185 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 244
Db 190 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 249
Qy 245 VPKKP 249

```

```

Db 250 VPKKP 254

```

RESULT 5

```

US-10-131-409-92
; Sequence 92, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
; APPLICANT: Malyankar et al.
; TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675C1P1CON1
; CURRENT FILING DATE: 2002-10-24
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-409-92

```

```

Query Match 98.2%; Score 1368; DB 14; Length 256;
Best Local Similarity 99.6%; Pred. No. 4.3e-122;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 IYGGGEAERSKMPQVSLRKHVGRPYWMMHFCGSSLIHPQWVLTAAACVGPVDVRLAALRVQL 64
Db 12 IYGGGEAERSKMPQVSLRKHVGRPYWMMHFCGSSLIHPQWVLTAAACVGPVDVRLAALRVQL 71
Qy 65 RBOHLYYDDQLLPPRSRIIVHPQFYTAQIGADIALLELEBPVKVSSHVTITLPPASSTFP 124
Db 72 RBOHLYYDDQLLPPRSRIIVHPQFYTAQIGADIALLELEBPVKVSSHVTITLPPASSTFP 131
Qy 125 PGMPCWVTGMDVNDERLPPFPPLKQVYPIIMENHI CDAKYHLGAYTGDVRIYRDDML 184
Db 132 PGMPCWVTGMDVNDERLPPFPPLKQVYPIIMENHI CDAKYHLGAYTGDVRIYRDDML 191
Qy 185 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 244
Db 192 CAGNTRRSDSCGDSGGLVCKVNGTWMLOAGVVSWSGCGAQNRPKGIYTRVYYLDMIHNY 251
Qy 245 VPKKP 249
Db 252 VPKKP 256

```

RESULT 6

```

US-10-139-854-92
; Sequence 92, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kumud

```

```

; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CON2
; CURRENT APPLICATION NUMBER: US/10/139,854
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 09/783,429
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; PRIOR FILING DATE: 2000-09-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-139-854-92

```

```

Query Match          98.2%; Score 1368; DB 15; Length 256;
Best Local Similarity 99.6%; Pred. No. 4,3e-122;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYGGQEARSKMFWQVSLRVHGPRYMMHFCGSLIHPQWLTAAACVGPVDKDLAALRYQL 64
DB 12 IYGGQEARSKMFWQVSLRVHGPRYMMHFCGSLIHPQWLTAAACVGPVDKDLAALRYQL 71
QY 65 REQHLYYDDQLLPVSRITIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASETFP 124
DB 72 REQHLYYDDQLLPVSRITIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASETFP 131
QY 125 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKYHLGAYTGDDVRIYRDML 184
DB 132 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKYHLGAYTGDDVRIYRDML 191
QY 185 CAGNTRRDSGQSDSGPRLVCVKNGTWLDAGVVSWSGEGCAQPNRPRTIYRVTYYLDMIHNY 244
DB 192 CAGNTRRDSGQSDSGPRLVCVKNGTWLDAGVVSWSGEGCAQPNRPRTIYRVTYYLDMIHNY 251
QY 245 VPKKP 249
DB 252 VPKKP 256

```

```

; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; PRIOR FILING DATE: 2000-09-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-150-813-92

```

```

Query Match          98.2%; Score 1368; DB 15; Length 256;
Best Local Similarity 99.6%; Pred. No. 4,3e-122;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYGGQEARSKMFWQVSLRVHGPRYMMHFCGSLIHPQWLTAAACVGPVDKDLAALRYQL 64
DB 12 IYGGQEARSKMFWQVSLRVHGPRYMMHFCGSLIHPQWLTAAACVGPVDKDLAALRYQL 71
QY 65 REQHLYYDDQLLPVSRITIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASETFP 124
DB 72 REQHLYYDDQLLPVSRITIVHPQFYTAQIGADIALLELEBPVKVSSHVHTVTLPPASETFP 131
QY 125 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKYHLGAYTGDDVRIYRDML 184
DB 132 PGMPCWVTGKGDVNDERLPPFPPLKQVKVPIEMENHICDAKYHLGAYTGDDVRIYRDML 191
QY 185 CAGNTRRDSGQSDSGPRLVCVKNGTWLDAGVVSWSGEGCAQPNRPRTIYRVTYYLDMIHNY 244
DB 192 CAGNTRRDSGQSDSGPRLVCVKNGTWLDAGVVSWSGEGCAQPNRPRTIYRVTYYLDMIHNY 251
QY 245 VPKKP 249
DB 252 VPKKP 256

```

```

RESULT 8
US-10-150-811-92
; Sequence 92, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
; APPLICANT: Malyankar et al.
; TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/150,811
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 09/970,607
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23

```


PRIOR APPLICATION NUMBER: 60/184,744
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: 60/197,083
 PRIOR FILING DATE: 2000-04-13
 PRIOR APPLICATION NUMBER: 60/233,405
 PRIOR FILING DATE: 2000-09-18
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 138
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 92
 LENGTH: 256
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-150-811-92

Query Match 98.2%; Score 1368; DB 15; Length 256;
 Best Local Similarity 99.6%; Pred. No. 4.3e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYGGGAPRSKMPQVSLRVHGPYMMHFCGSSLIHPQWLTAAACVGPDKDLAALRYQL 64
 DB 12 IYGGGAPRSKMPQVSLRVHGPYMMHFCGSSLIHPQWLTAAACVGPDKDLAALRYQL 71
 QY 65 RQHLHYDQDLLPVSRIIVHPQFYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 124
 DB 72 RQHLHYDQDLLPVSRIIVHPQFYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 131
 QY 125 PGMPCWVTGMGDVNDERLPPPPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 184
 DB 132 PGMPCWVTGMGDVNDERLPPPPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 191
 QY 185 CAGNTRRDSGCGSGPLVCKVNGTWTLAGVYVSWGCGAQNRRPGIYTRVYYLDMIHYY 244
 DB 192 CAGNTRRDSGCGSGPLVCKVNGTWTLAGVYVSWGCGAQNRRPGIYTRVYYLDMIHYY 251
 QY 245 VPKKP 249
 DB 252 VPKKP 256

RESULT 9
 US-10-131-409-104
 ; Sequence 104, Application US/10131409
 ; Publication No. US20030199465A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Malyanar et al.
 ; TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Sam
 ; FILE REFERENCE: 15966-675CPICON1
 ; CURRENT APPLICATION NUMBER: US/10/131,409
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/898,954
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 104
 LENGTH: 264
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-131-409-104

Query Match 98.2%; Score 1368; DB 14; Length 264;
 Best Local Similarity 99.6%; Pred. No. 4.4e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYGGGAPRSKMPQVSLRVHGPYMMHFCGSSLIHPQWLTAAACVGPDKDLAALRYQL 64
 DB 20 IYGGGAPRSKMPQVSLRVHGPYMMHFCGSSLIHPQWLTAAACVGPDKDLAALRYQL 79
 QY 65 RQHLHYDQDLLPVSRIIVHPQFYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 124
 DB 80 RQHLHYDQDLLPVSRIIVHPQFYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 139
 QY 125 PGMPCWVTGMGDVNDERLPPPPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 184
 DB 140 PGMPCWVTGMGDVNDERLPPPPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIYRDDML 199
 QY 185 CAGNTRRDSGCGSGPLVCKVNGTWTLAGVYVSWGCGAQNRRPGIYTRVYYLDMIHYY 244
 DB 200 CAGNTRRDSGCGSGPLVCKVNGTWTLAGVYVSWGCGAQNRRPGIYTRVYYLDMIHYY 259
 QY 245 VPKKP 249
 DB 260 VPKKP 264

RESULT 10
 US-10-139-854-104
 ; Sequence 104, Application US/10139854
 ; Publication No. US20030202971A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder, Kumud
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-675CON2
 ; CURRENT APPLICATION NUMBER: US/10/139,854
 ; CURRENT FILING DATE: 2002-12-02
 ; PRIOR APPLICATION NUMBER: 09/783,429
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 104
 ; LENGTH: 264
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-139-854-104

Query Match 98.2%; Score 1368; DB 15; Length 264;
 Best Local Similarity 99.6%; Pred. No. 4.4e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 64
 DB 20 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 79
 QY 65 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 124
 DB 80 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 139
 QY 125 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 DB 140 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 199
 QY 185 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 244
 DB 200 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 259
 QY 245 VPKKP 249
 DB 260 VPKKP 264

RESULT 11

US-10-150-813-104
 ; Sequence 104, Application US/10150813
 ; Publication No. US20030224367A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MalyanKar
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-675CIP2CON2
 ; CURRENT APPLICATION NUMBER: US/10/150,813
 ; CURRENT FILING DATE: 2002-12-03
 ; PRIOR APPLICATION NUMBER: 09/898,954
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 104
 ; LENGTH: 264
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-150-813-104

Query Match 98.2%; Score 1368; DB 15; Length 264;
 Best Local Similarity 99.6%; Pred. No. 4,4e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 64
 DB 20 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 79
 QY 65 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 124
 DB 80 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 139

QY 125 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 DB 140 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 199
 QY 185 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 244
 DB 200 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 259
 QY 245 VPKKP 249
 DB 260 VPKKP 264

RESULT 12

US-10-150-811-104
 ; Sequence 104, Application US/10150811
 ; Publication No. US20040010120A1e1 Polypeptides and Nucleic Acids Encoding Same
 ; GENERAL INFORMATION:
 ; APPLICANT: MalyanKar et al.
 ; TITLE OF INVENTION: No. US20040010120A1e1 Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-675CIP2CON1
 ; CURRENT APPLICATION NUMBER: US/10/150,811
 ; CURRENT FILING DATE: 2002-05-17
 ; PRIOR APPLICATION NUMBER: 09/970,607
 ; PRIOR FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: 60/182,733
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/182,724
 ; PRIOR FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 60/183,896
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: 60/184,497
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/224,157
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: 60/184,482
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,744
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/197,083
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 60/233,405
 ; PRIOR FILING DATE: 2000-09-18
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 138
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 104
 ; LENGTH: 264
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-150-811-104

Query Match 98.2%; Score 1368; DB 15; Length 264;
 Best Local Similarity 99.6%; Pred. No. 4,4e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 64
 DB 20 IYVGGQEARPSKWPQVSLRVHNGPYMMHFCGSSLIHPQWVLTAAACVGPDKDLAALRVQL 79
 QY 65 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 124
 DB 80 REQHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLELEBEPVKVSSHVHTVTLPPASSTFP 139
 QY 125 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 DB 140 PGMPCWVTGMDVNDERLPPFPPLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 199
 QY 185 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 244
 DB 200 CAGNTRRDSCGDSGGPVLCKVNGTWTLAGVSVWGECAQPNRPGIYTRVYLLDWHIHHY 259
 QY 245 VPKKP 249

Db 260 VPKKP 264

RESULT 13
 US-09-813-432-44
 ; Sequence 44, Application US/09813432
 ; Publication No. US20030148485A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Taupier Jr., Raymond J
 ; APPLICANT: Majmuder, Kamid
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Mezei, Peter S
 ; APPLICANT: Verneq, Corine A. M.
 ; TITLE OF INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same
 ; FILE REFERENCE: 15966-729
 ; CURRENT APPLICATION NUMBER: US/09/813,432
 ; PRIORITY FILING DATE: 2001-03-20
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,835
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,768
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,972
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,199
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,947
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 60/192,665
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,657
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,984
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,664
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,836
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/193,843
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 44
 ; LENGTH: 275
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-813-432-44

Query Match 98.2%; Score 1368; DB 10; Length 275;
 Best Local Similarity 99.6%; Pred. No. 4.7e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOVLTAAACVGPVDVLDLAALRVQL 64
 |||||
 Db 31 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOVLTAAACVGPVDVLDLAALRVQL 90
 |||||
 QY 65 RBOHLYYODQLLPPSRRIIVHPOFYTAOIGADIALLELEBEPVKSVMHTVTLPPASETFP 124
 |||||
 Db 91 RBOHLYYODQLLPPSRRIIVHPOFYTAOIGADIALLELEBEPVKSVMHTVTLPPASETFP 150
 |||||
 QY 125 PMPFCWVTGMDVNDERLPPRPPLKQVAVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 |||||
 Db 151 PMPFCWVTGMDVNDERLPPRPPLKQVAVPIMENHICDAKXHLGAYTGDDVRIVRDDML 210
 |||||
 QY 185 CAGNTRRSDSCGDSGGPLVCKVNGTWTLAGVVSWEBCGCAQPNRPGIYTRVTVYYLDWIHHY 244
 |||||
 Db 211 CAGNTRRSDSCGDSGGPLVCKVNGTWTLAGVVSWEBCGCAQPNRPGIYTRVTVYYLDWIHHY 270
 |||||
 QY 245 VPKKP 249
 |||||
 Db 271 VPKKP 275

Db 260 VPKKP 264

RESULT 14
 US-10-352-684A-48
 ; Sequence 48, Application US/10352684A
 ; Publication No. US20030215452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals Inc.
 ; APPLICANT: Carroll, Joseph M.
 ; APPLICANT: Healy, Aileen
 ; APPLICANT: Welch, Nadine S.
 ; APPLICANT: Kelly, Louise M.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
 ; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
 ; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
 ; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
 ; FILE REFERENCE: MP102-019P1RNONMIM
 ; CURRENT APPLICATION NUMBER: US/10/352,684A
 ; PRIORITY FILING DATE: 2003-01-28
 ; PRIOR FILING DATE: 2003-01-28
 ; PRIOR APPLICATION NUMBER: US 60/354,333
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: US 60/360,258
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/364,476
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: US 60/375,626
 ; PRIOR FILING DATE: 2002-04-26
 ; PRIOR APPLICATION NUMBER: US 60/386,494
 ; PRIOR FILING DATE: 2002-06-06
 ; PRIOR APPLICATION NUMBER: US 60/390,965
 ; PRIOR FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: US 60/392,480
 ; PRIOR FILING DATE: 2002-06-28
 ; PRIOR APPLICATION NUMBER: US 60/394,128
 ; PRIOR FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: US 60/399,783
 ; PRIOR FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: US 60/403,221
 ; PRIOR FILING DATE: 2002-08-13
 ; Remaining Prior Application data removed - See file wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 48
 ; LENGTH: 275
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-352-684A-48

Query Match 98.2%; Score 1368; DB 15; Length 275;
 Best Local Similarity 99.6%; Pred. No. 4.7e-122;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOVLTAAACVGPVDVLDLAALRVQL 64
 |||||
 Db 31 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOVLTAAACVGPVDVLDLAALRVQL 90
 |||||
 QY 65 RBOHLYYODQLLPPSRRIIVHPOFYTAOIGADIALLELEBEPVKSVMHTVTLPPASETFP 124
 |||||
 Db 91 RBOHLYYODQLLPPSRRIIVHPOFYTAOIGADIALLELEBEPVKSVMHTVTLPPASETFP 150
 |||||
 QY 125 PMPFCWVTGMDVNDERLPPRPPLKQVAVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 |||||
 Db 151 PMPFCWVTGMDVNDERLPPRPPLKQVAVPIMENHICDAKXHLGAYTGDDVRIVRDDML 210
 |||||
 QY 185 CAGNTRRSDSCGDSGGPLVCKVNGTWTLAGVVSWEBCGCAQPNRPGIYTRVTVYYLDWIHHY 244
 |||||
 Db 211 CAGNTRRSDSCGDSGGPLVCKVNGTWTLAGVVSWEBCGCAQPNRPGIYTRVTVYYLDWIHHY 270
 |||||
 QY 245 VPKKP 249
 |||||
 Db 271 VPKKP 275

```

RESULT 15
US-10-174-364-44      ; Sequence 44, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729C1P2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 44
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-364-44

```

```

Query Match          98.2%; Score 1368; DB 15; Length 275;
Best Local Similarity 99.6%; Pred. No. 4,7e-122;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IYGGDAPRSKMPQVSLRVNRPYMNFCGSLINHPQWLTAAACVGRDVKDIALALRVQL 64
DB 31 IYGGDAPRSKMPQVSLRVNRPYMNFCGSLINHPQWLTAAACVGRDVKDIALALRVQL 50
QY 65 REQHLVYDQDLIPVSRITVHNRQFYTAQIGADIALLELEBPVKVSSHVHTVTLRPASETPR 124
DB 91 REQHLVYDQDLIPVSRITVHNRQFYTAQIGADIALLELEBPVKVSSHVHTVTLRPASETPR 150
QY 125 PGNPCWVTGWDVNDERLRRPRLKQVKVPIEMENHICDAKYNHGAUTGDDVRIYRDDML 184
DB 151 PGNPCWVTGWDVNDERLRRPRLKQVKVPIEMENHICDAKYNHGAUTGDDVRIYRDDML 210
QY 185 CAGNTRRDSGGSPGGPVLVCKVNGTMDQAGVVSWSGSCAQPNNRPGIYTRVTVYLLDMIHNY 244
DB 211 CAGNTRRDSGGSPGGPVLVCKVNGTMDQAGVVSWSGSCAQPNNRPGIYTRVTVYLLDMIHNY 270
QY 245 VPKKP 249
DB 271 VPKKP 275

```

Search completed: August 27, 2005, 23:24:28
Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 21:22:59 ; Search time 62 Seconds
(without alignments)
299.800 Million cell updates/sec

Title: US-09-598-982C-21
Perfect score: 1393
Sequence: 1 LEKRIVGGQEARPRSKMPQV.....IYTRVTVYLDWIMHHVPPKKP 249

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1387	99.6	249	3	US-09-079-970A-5
2	1368	98.2	245	3	US-09-079-970A-6
3	1368	98.2	245	4	US-09-601-318-1
4	1368	98.2	274	2	US-09-016-366A-21
5	1368	98.2	274	2	US-08-978-404B-16
6	1363	97.8	273	2	US-09-016-366A-19
7	1363	97.8	273	2	US-08-978-404B-14
8	1361	97.7	244	4	US-09-601-318-4
9	1361	97.7	244	4	US-09-601-318-5
10	1361	97.7	244	4	US-09-601-318-6
11	1361	97.7	244	4	US-09-601-318-7
12	1344	96.5	267	2	US-09-016-366A-23
13	1344	96.5	267	2	US-08-978-404B-18
14	1344	96.5	267	4	US-09-917-254-101
15	1274	91.5	245	3	US-08-944-483-69
16	1274	91.5	275	2	US-09-016-366A-17
17	1274	91.5	275	2	US-08-978-404B-12
18	1091	78.3	276	2	US-09-016-366A-15
19	1091	78.3	276	2	US-08-978-404B-21
20	1075	77.2	270	2	US-08-978-404B-8
21	1054	75.7	273	2	US-08-978-404B-3
22	1051	75.4	274	2	US-08-978-404B-5
23	1037	74.4	273	2	US-08-978-404B-6
24	702	50.4	190	3	US-08-845-998-4
25	702	50.4	190	3	US-09-206-537-4
26	702	50.4	190	3	US-09-430-854-4
27	698	50.1	190	2	US-08-845-998-6

ALIGNMENTS

Query Match	Score	DB 3j	Length	DB	Sequence
99.6%	1387	3j	249	3	US-09-206-537-6
99.6%	1368	3j	245	3	US-09-430-854-6
99.6%	1368	3j	245	4	US-08-978-404B-10
99.6%	1368	3j	274	2	US-09-386-653A-7
99.6%	1368	3j	274	2	US-09-386-653A-9
99.6%	1363	3j	273	2	US-09-387-375-7
99.6%	1363	3j	273	2	US-10-041-400A-7
99.6%	1363	3j	273	2	US-10-042-091A-7
99.6%	1361	3j	244	4	US-09-387-375-9
99.6%	1361	3j	244	4	US-10-041-400A-9
99.6%	1361	3j	244	4	US-10-042-091A-9
99.6%	1361	3j	244	4	US-09-386-642-12
99.6%	1361	3j	244	4	US-09-386-642-11
99.6%	1361	3j	244	4	US-08-944-483-66
99.6%	1361	3j	244	4	US-09-386-629-7
99.6%	1361	3j	244	4	US-09-907-794A-263
99.6%	1361	3j	244	4	US-09-905-125A-263
99.6%	1361	3j	244	4	US-09-902-775A-263

RESULT 1
US-09-079-970A-5
Sequence 5, Application US/09079970A
Patent No. 6274366
GENERAL INFORMATION:
APPLICANT: Mafflet, Mark A.
APPLICANT: Niles, Andrew L.
TITLE OF INVENTION: Enzymatically-Active Recombinant Human
TITLE OF INVENTION: Beta-Trypsin and Method of Making Same
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Intellectual Property Department
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WISCONSIN
COUNTRY: U.S.A.
ZIP: 53717-1914.
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079, 970A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leone, Joseph T.
REGISTRATION NUMBER: 37,170
REFERENCE/DOCKET NUMBER: 34506.073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2108
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-079-970A-5

Query Match 99.6%; Score 1387; DB 3j; Length 249;
Best Local Similarity 99.6%; Pred. No. 7.98-141;
Matches 248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LEKRIVGGQEARPRSKMPQVSLRVHGPYRMMHFCGSLIHPQWVITLAAACVGPDYKDLAAL 60
DB 1 LEKRIVGGQEARPRSKMPQVSLRVHGPYRMMHFCGSLIHPQWVITLAAACVGPDYKDLAAL 60

QY 61 RVLQREOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPAS 120
 Db 61 RVLQREOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPAS 120
 QY 121 ETRPPGKPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYR 180
 Db 121 ETRPPGKPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYR 180
 QY 181 DDMLCAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM 240
 Db 181 DDMLCAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM 240
 QY 241 IHHYVPRKP 249
 Db 241 IHHYVPRKP 249

RESULT 2

US-09-979-970A-6
 ; Sequence 6, Application US/09079970A
 ; Patent No. 6274366
 ; GENERAL INFORMATION:
 ; APPLICANT: Maffitt, Mark A.
 ; APPLICANT: Niles, Andrew L.
 ; APPLICANT: Haak-Frendscho, Mary
 ; TITLE OF INVENTION: Enzymatically-Active Recombinant Human
 ; TITLE OF INVENTION: Beta-Trypsin and Method of Making Same
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Intellectual Property Department
 ; STREET: 8000 Excelsior Drive, Suite 401
 ; CITY: Madison
 ; STATE: WISCONSIN
 ; COUNTRY: U.S.A.
 ; ZIP: 53717-1914
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/079,970A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Leone, Joseph T.
 ; REGISTRATION NUMBER: 37,170
 ; REFERENCE/DOCKET NUMBER: 34506.073
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 831-2100
 ; TELEFAX: (608) 831-2106
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 245 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-079-970A-6

Query Match 98.2%; Score 1368; DB 3; Length 245;
 Best Local Similarity 99.6%; Pred. No. 8.4e-139;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPSKWPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVYKDLAALRYQL 64
 Db 1 IVGGQEARPSKWPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVYKDLAALRYQL 60
 QY 65 REOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPASSTFP 124
 Db 61 REOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPASSTFP 120
 QY 125 PGMPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYRDDML 184
 Db 125 PGMPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYRDDML 184

Db 121 PGMPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYRDDML 180
 QY 185 CAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM IHY 244
 Db 181 CAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM IHY 240
 QY 245 VPKKP 249
 Db 241 VPKKP 245

RESULT 3

US-09-601-318-1
 ; Sequence 1, Application US/09601318
 ; Patent No. 6613769
 ; GENERAL INFORMATION:
 ; APPLICANT: Max-Planck-Gesellschaft z. Fnd. d. Wissensch. e.V
 ; APPLICANT: Byk Gulden Lomborg Chemische Fabrik GmbH
 ; APPLICANT: Bode, Wolfram
 ; APPLICANT: Moroder, Luis
 ; APPLICANT: Pereira, Pedro Jose Barbosa
 ; APPLICANT: Bergner, Andreas
 ; APPLICANT: Huber, Robert
 ; APPLICANT: Sommerhoff, Christian
 ; APPLICANT: Schaschke, No. 6613769dbert
 ; APPLICANT: Br, Thomas
 ; APPLICANT: Martin, Thomas
 ; APPLICANT: Stadlwieser, Josef
 ; APPLICANT: Ulrich, Wolf-Rdiger
 ; APPLICANT: Dominik, Andreas
 ; APPLICANT: Thibaut, Ulrich
 ; APPLICANT: Bundschuh, Daniela
 ; APPLICANT: Beume, Rolf
 ; APPLICANT: Goebel, Karl-Josef
 ; TITLE OF INVENTION: Trypsin-Inhibitor
 ; FILE REFERENCE: 17674P WO-1
 ; CURRENT APPLICATION NUMBER: US/09/601,318
 ; PRIOR FILING DATE: 2001-01-22
 ; PRIOR APPLICATION NUMBER: DE19804761.4
 ; PRIOR FILING DATE: 1998-02-06
 ; PRIOR APPLICATION NUMBER: DE19851300.3
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 245
 ; TYPR: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-601-318-1

Query Match 98.2%; Score 1368; DB 4; Length 245;
 Best Local Similarity 99.6%; Pred. No. 8.4e-139;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPSKWPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVYKDLAALRYQL 64
 Db 1 IVGGQEARPSKWPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVYKDLAALRYQL 60
 QY 65 REOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPASSTFP 124
 Db 61 REOHLYYQDQLLPSVRIIVHPQFYTAQIGADIALLEBEPVYVSSHVHTVTLPPASSTFP 120
 QY 125 PGMPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYRDDML 184
 Db 121 PGMPCWVTGMDVNDERLPPRPLKQYKVPIMENHICDAKYNLGAAYTGDVRIYRDDML 180
 QY 185 CAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM IHY 244
 Db 181 CAGNTRRDSGQDSSGSPLVCKVNGTMDQAGVYVSWGEGCAQPNRPGIYTRVYYLDM IHY 240
 QY 245 VPKKP 249
 Db 241 VPKKP 245

RESULT 4
US-09-016-366A-21

; Sequence 21, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-016-366A-21

Query Match 98.2%; Score 1368; DB 2; Length 274;
Best Local Similarity 99.6%; Pred. No. 9.8e-139;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLAAALRVQL 64
DB 30 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLAAALRVQL 89
QY 65 REQHLYYQDQLLPSRRIIVHQPFTYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 124
DB 90 REQHLYYQDQLLPSRRIIVHQPFTYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 149
QY 125 PGMPCWVTGMDVNDDELPPFPPLKQVVPIMENHI CDACKYHGAATYGDVRIVRDDML 184
DB 150 PGMPCWVTGMDVNDDELPPFPPLKQVVPIMENHI CDACKYHGAATYGDVRIVRDDML 209
QY 185 CAGNTRRDS CGDSDGGLVCKVNGTMTLQAGVVSMBEGCAQPNRPGIYTRVTVYLDWIHHY 244
DB 210 CAGNTRRDS CGDSDGGLVCKVNGTMTLQAGVVSMBEGCAQPNRPGIYTRVTVYLDWIHHY 269
QY 245 VPKKP 249
DB 270 VPKKP 274

US-08-978-404B-16

; Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5968782e
; US-08-978-404B-16

Query Match 98.2%; Score 1368; DB 2; Length 274;
Best Local Similarity 99.6%; Pred. No. 9.8e-139;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLAAALRVQL 64
DB 30 IVGGGEAPRSKMPQVSLRVHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLAAALRVQL 89
QY 65 REQHLYYQDQLLPSRRIIVHQPFTYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 124
DB 90 REQHLYYQDQLLPSRRIIVHQPFTYTAQIGADIALLEBEPVKVSSHVHTVTLPPASSTFP 149
QY 125 PGMPCWVTGMDVNDDELPPFPPLKQVVPIMENHI CDACKYHGAATYGDVRIVRDDML 184
DB 150 PGMPCWVTGMDVNDDELPPFPPLKQVVPIMENHI CDACKYHGAATYGDVRIVRDDML 209
QY 185 CAGNTRRDS CGDSDGGLVCKVNGTMTLQAGVVSMBEGCAQPNRPGIYTRVTVYLDWIHHY 244
DB 210 CAGNTRRDS CGDSDGGLVCKVNGTMTLQAGVVSMBEGCAQPNRPGIYTRVTVYLDWIHHY 269
QY 245 VPKKP 249
DB 270 VPKKP 274

RESULT 5

RESULT 6

```

; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-09-016-366A-19
;
; Query Match          97.8%; Score 1363; DB 2; Length 273;
; Best Local Similarity 99.2%; Pred. No. 3.4e-138;
; Matches 243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 5 IVGGGEARSRKRWQVSLRVHGRVYMHFCGSSLIHPOVWLTAAACVGRVVKDLAALRVQL 64
;   |||
; Db 29 IVGGGEARSRKRWQVSLRVHGRVYMHFCGSSLIHPOVWLTAAACVGRVVKDLAALRVQL 88
;
; QY 65 REQHLYYDQDLPLVSRILVHPOFYTAQIGADIALLELEBEPVYSSHVHTVTLPPASETFP 124
;   |||
; Db 89 REQHLYYDQDLPLVSRILVHPOFYTAQIGADIALLELEBEPVYSSHVHTVTLPPASETFP 148
;
; QY 125 PGMPCWVTGWDVNDERLPPRFPKQVKVPIEMENHICDAKYHLAGAYTGDVRIYRDMWL 184
;   |||
; Db 149 PGMPCWVTGWDVNDERLPPRFPKQVKVPIEMENHICDAKYHLAGAYTGDVRIYRDMWL 208
;
; QY 185 CAGNTRRSDSCGDSGGPLVCKVNGTWLQAGVYVSWGEGCAQPNRPGIYTRVYIYLDWIMHHY 244
;   |||
; Db 209 CAGNTRRSDSCGDSGGPLVCKVNGTWLQAGVYVSWGEGCAQPNRPGIYTRVYIYLDWIMHHY 268
;
; QY 245 VPKKP 249
;   |||
; Db 269 VPKKP 273
;
; RESULT 7
; US-08-978-404B-14
; Sequence 14, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES

```

```

; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5968782e
;
; US-08-978-404B-14
;
; Query Match          97.8%; Score 1363; DB 2; Length 273;
; Best Local Similarity 99.2%; Pred. No. 3.4e-138;
; Matches 243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 5 IVGGGEARSRKRWQVSLRVHGRVYMHFCGSSLIHPOVWLTAAACVGRVVKDLAALRVQL 64
;   |||
; Db 29 IVGGGEARSRKRWQVSLRVHGRVYMHFCGSSLIHPOVWLTAAACVGRVVKDLAALRVQL 88
;
; QY 65 REQHLYYDQDLPLVSRILVHPOFYTAQIGADIALLELEBEPVYSSHVHTVTLPPASETFP 124
;   |||
; Db 89 REQHLYYDQDLPLVSRILVHPOFYTAQIGADIALLELEBEPVYSSHVHTVTLPPASETFP 148
;
; QY 125 PGMPCWVTGWDVNDERLPPRFPKQVKVPIEMENHICDAKYHLAGAYTGDVRIYRDMWL 184
;   |||
; Db 149 PGMPCWVTGWDVNDERLPPRFPKQVKVPIEMENHICDAKYHLAGAYTGDVRIYRDMWL 208
;
; QY 185 CAGNTRRSDSCGDSGGPLVCKVNGTWLQAGVYVSWGEGCAQPNRPGIYTRVYIYLDWIMHHY 244
;   |||
; Db 209 CAGNTRRSDSCGDSGGPLVCKVNGTWLQAGVYVSWGEGCAQPNRPGIYTRVYIYLDWIMHHY 268
;
; QY 245 VPKKP 249
;   |||
; Db 269 VPKKP 273
;
; RESULT 8
; US-09-601-318-4
; Sequence 4, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft z. Fnd. d. Wissensch. e.V
; APPLICANT: Byk Gulden Lomborg Chemische Fabrik GmbH
; APPLICANT: Bode, Wolfram
; APPLICANT: Moroder, Luis
; APPLICANT: Pereira, Pedro Jose Barbosa

```



```

; APPLICANT: Bergner, Andreas
; APPLICANT: Huber, Robert
; APPLICANT: Sommerhoff, Christian
; APPLICANT: Schaschke, No. 6613769bert
; APPLICANT: Br, Thomas
; APPLICANT: Martin, Thomas
; APPLICANT: Stadlwieser, Josef
; APPLICANT: Ulrich, Wolf-Rdiger
; APPLICANT: Dominik, Andreas
; APPLICANT: Thibaut, Ulrich
; APPLICANT: Bundschuh, Daniela
; APPLICANT: Beume, Rolf
; APPLICANT: Goebel, Karl-Josef
; TITLE OF INVENTION: Tryptase-Inhibitoren
; FILE REFERENCE: 1/674P WO-1
; CURRENT APPLICATION NUMBER: US/09/601,318
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: DE19804761.4
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: DE19851300.3
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-601-318-4

```

```

Query Match 97.7%; Score 1361; DB 4; Length 244;
Best Local Similarity 99.6%; Pred. No. 4,7e-138;
Matches 243; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 5 IVGGQEARPKRMPQVSLRVRHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLALRYQL 64
DB 1 IVGGQEARPKRMPQVSLRVRHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLALRYQL 60
QY 65 REQHLYYQDQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVHTVTLPPASSTFP 124
DB 61 REQHLYYQDQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVHTVTLPPASSTFP 120
QY 125 PGMPCWVTGMDVNDRLRPPRPLKQVYRIMENHICDAKYHLGAYTGDDVRIVRDML 184
DB 121 PGMPCWVTGMDVNDRLRPPRPLKQVYRIMENHICDAKYHLGAYTGDDVRIVRDML 180
QY 185 CAGNTRRDSGCGDGGPVLCKVNGTWTWLDAGVVSMBEGCAQPNRPGIYTRVYTYLDWIHHY 244
DB 181 CAGNTRRDSGCGDGGPVLCKVNGTWTWLDAGVVSMBEGCAQPNRPGIYTRVYTYLDWIHHY 240
QY 245 VPKK 248
DB 241 VPKK 244

```

```

; APPLICANT: Thibaut, Ulrich
; APPLICANT: Bundschuh, Daniela
; APPLICANT: Beume, Rolf
; APPLICANT: Goebel, Karl-Josef
; TITLE OF INVENTION: Tryptase-Inhibitoren
; FILE REFERENCE: 1/674P WO-1
; CURRENT APPLICATION NUMBER: US/09/601,318
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: DE19804761.4
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: DE19851300.3
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-601-318-5

```

```

Query Match 97.7%; Score 1361; DB 4; Length 244;
Best Local Similarity 99.6%; Pred. No. 4,7e-138;
Matches 243; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 5 IVGGQEARPKRMPQVSLRVRHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLALRYQL 64
DB 1 IVGGQEARPKRMPQVSLRVRHGRPYMMHFCGSSLIHPOWVLTAAACVGPVDVLDLALRYQL 60
QY 65 REQHLYYQDQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVHTVTLPPASSTFP 124
DB 61 REQHLYYQDQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVHTVTLPPASSTFP 120
QY 125 PGMPCWVTGMDVNDRLRPPRPLKQVYRIMENHICDAKYHLGAYTGDDVRIVRDML 184
DB 121 PGMPCWVTGMDVNDRLRPPRPLKQVYRIMENHICDAKYHLGAYTGDDVRIVRDML 180
QY 185 CAGNTRRDSGCGDGGPVLCKVNGTWTWLDAGVVSMBEGCAQPNRPGIYTRVYTYLDWIHHY 244
DB 181 CAGNTRRDSGCGDGGPVLCKVNGTWTWLDAGVVSMBEGCAQPNRPGIYTRVYTYLDWIHHY 240
QY 245 VPKK 248
DB 241 VPKK 244

```

```

RESULT 10
US-09-601-318-6
; Sequence 6, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft z. Fnd. d. Wissensch. e.V
; APPLICANT: Sommerhoff, Christian
; APPLICANT: Br, Thomas
; APPLICANT: Schaschke, No. 6613769bert
; APPLICANT: Martin, Thomas
; APPLICANT: Stadlwieser, Josef
; APPLICANT: Ulrich, Wolf-Rdiger
; APPLICANT: Dominik, Andreas
; APPLICANT: Thibaut, Ulrich
; APPLICANT: Bundschuh, Daniela
; APPLICANT: Beume, Rolf
; APPLICANT: Goebel, Karl-Josef
; TITLE OF INVENTION: Tryptase-Inhibitoren
; FILE REFERENCE: 1/674P WO-1
; CURRENT APPLICATION NUMBER: US/09/601,318
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: DE19804761.4

```

/ PRIOR FILING DATE: 1998-02-06
 / PRIOR APPLICATION NUMBER: DE19851300.3
 / PRIOR FILING DATE: 1998-11-06
 / NUMBER OF SEQ ID NOS: 7
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 6
 / LENGTH: 244
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-601-318-6

Query Match 97.7%; Score 1361; DB 4; Length 244;
 Best Local Similarity 99.6%; Pred. No. 4.7e-138;
 Matches 243; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPKWQVSLRHHGPRYMMHFCGSSLHHPQWVLTAAACVGPDKDIALRVQL 64
 1 IVGGQEARPKWQVSLRHHGPRYMMHFCGSSLHHPQWVLTAAACVGPDKDIALRVQL 60
 QY 65 REQHLYYDQDLFVSRITVHPQFYTAQIGADIALLEBEPVSSHVHTVTLPPASETFP 124
 61 REQHLYYDQDLFVSRITVHPQFYTAQIGADIALLEBEPVSSHVHTVTLPPASETFP 120
 QY 125 PGMPCWVTGMDVNDERLPPFPRLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 121 PGMPCWVTGMDVNDERLPPFPRLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 180
 QY 185 CAGNTRRDSGQDSSGPIVCKVNGTWLQAGVSWGEGCAQPNRPGIYTRVYYLDWIHHY 244
 181 CAGNTRRDSGQDSSGPIVCKVNGTWLQAGVSWGEGCAQPNRPGIYTRVYYLDWIHHY 240
 QY 245 VPKK 248
 241 VPKK 244

RESULT 11

/ US-09-601-318-7
 / Sequence 7, Application US/09601318
 / Patent No. 6613769
 / GENERAL INFORMATION:
 / APPLICANT: Byk Gulden Lomberg Chemische Fabrik GmbH
 / APPLICANT: Bode, Wolfram
 / APPLICANT: Moroder, Luis
 / APPLICANT: Pereira, Pedro Jose Barbosa
 / APPLICANT: Bergner, Andreas
 / APPLICANT: Huber, Robert
 / APPLICANT: Sommerhoff, Christian
 / APPLICANT: Schaschke, No. 6613769bert
 / APPLICANT: Br, Thomas
 / APPLICANT: Martin, Thomas
 / APPLICANT: Stadlmaier, Josef
 / APPLICANT: Ulrich, Wolf-Rüdiger
 / APPLICANT: Dominik, Andreas
 / APPLICANT: Thibaut, Ulrich
 / APPLICANT: Bundeschuh, Daniela
 / APPLICANT: Beume, Rolf
 / APPLICANT: Goedel, Karl-Josef
 / TITLE OF INVENTION: Trypase-Inhibitoren
 / FILE REFERENCE: 17674P WO-1
 / CURRENT APPLICATION NUMBER: US/09/601, 318
 / FILING DATE: 2001-01-22
 / PRIOR APPLICATION NUMBER: DE19804761.4
 / PRIOR FILING DATE: 1998-02-06
 / PRIOR APPLICATION NUMBER: DE19851300.3
 / PRIOR FILING DATE: 1998-11-06
 / NUMBER OF SEQ ID NOS: 7
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 7
 / LENGTH: 244
 / TYPE: PRT
 / ORGANISM: Homo sapiens

US-09-601-318-7
 Query Match 97.7%; Score 1361; DB 4; Length 244;
 Best Local Similarity 99.6%; Pred. No. 4.7e-138;
 Matches 243; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IVGGQEARPKWQVSLRHHGPRYMMHFCGSSLHHPQWVLTAAACVGPDKDIALRVQL 64
 1 IVGGQEARPKWQVSLRHHGPRYMMHFCGSSLHHPQWVLTAAACVGPDKDIALRVQL 60
 QY 65 REQHLYYDQDLFVSRITVHPQFYTAQIGADIALLEBEPVSSHVHTVTLPPASETFP 124
 61 REQHLYYDQDLFVSRITVHPQFYTAQIGADIALLEBEPVSSHVHTVTLPPASETFP 120
 QY 125 PGMPCWVTGMDVNDERLPPFPRLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 184
 121 PGMPCWVTGMDVNDERLPPFPRLKQVVPIMENHICDAKXHLGAYTGDDVRIVRDDML 180
 QY 185 CAGNTRRDSGQDSSGPIVCKVNGTWLQAGVSWGEGCAQPNRPGIYTRVYYLDWIHHY 244
 181 CAGNTRRDSGQDSSGPIVCKVNGTWLQAGVSWGEGCAQPNRPGIYTRVYYLDWIHHY 240
 QY 245 VPKK 248
 241 VPKK 244

RESULT 12

/ US-09-016-366A-23
 / Sequence 23, Application US/09016366A
 / Patent No. 5955431
 / GENERAL INFORMATION:
 / APPLICANT: Stevens, Richard L.
 / APPLICANT: Huang, Chifu
 / TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 / NUMBER OF SEQUENCES: 65
 / CORRESPONDENCE ADDRESS:
 / ADDRESSER: Wolf, Greenfield & Sacks, P.C.
 / STREET: 600 Atlantic Avenue
 / CITY: Boston
 / STATE: MA
 / COUNTRY: U.S.A.
 / ZIP: 02210-2211
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Diskette
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM:
 / SOFTWARE: FASTSEQ for Windows Version 2.0
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/016, 366A
 / FILING DATE: January 30, 1998
 / CLASSIFICATION: 530
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/037, 090
 / FILING DATE: 05-FEB-1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Plumer, Elizabeth R.
 / REGISTRATION NUMBER: 36, 637
 / REFERENCE/DOCKET NUMBER: B0801/7093
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 617-720-3500
 / TELEFAX: 617-720-2441
 / TELEK:
 / INFORMATION FOR SEQ ID NO: 23:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 267 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / US-09-016-366A-23
 Query Match 96.5%; Score 1344; DB 2; Length 267;

Best Local Similarity 98.4%; Pred. No. 3.6e-136;
 Matches 241; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 64
 |||||
 DB 23 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 82

QY 65 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 124
 |||||
 DB 83 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 142

QY 125 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 184
 |||||
 DB 143 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 202

QY 185 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 244
 |||||
 DB 203 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 262

QY 245 VPKKP 249
 |||||
 DB 263 VPKKP 267

RESULT 13
 US-08-978-404B-18
 ; Sequence 18, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard J.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEK:
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: No. 5968782e
 ; US-08-978-404B-18

Query Match 96.5%; Score 1344; DB 2; Length 267;
 Best Local Similarity 98.4%; Pred. No. 3.6e-136;
 Matches 241; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 64
 |||||
 DB 23 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 82

QY 65 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 124
 |||||
 DB 83 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 142

QY 125 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 184
 |||||
 DB 143 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 202

QY 185 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 244
 |||||
 DB 203 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 262

QY 245 VPKKP 249
 |||||
 DB 263 VPKKP 267

RESULT 14
 US-09-917-254-101
 ; Sequence 101, Application US/09917254
 ; Patent No. 6703204
 ; GENERAL INFORMATION:
 ; APPLICANT: Mutter, George
 ; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
 ; FILE REFERENCE: B0801/7224 (SRV)
 ; CURRENT APPLICATION NUMBER: US/09/917,254
 ; PRIORITY FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: US 60/222,093
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 101
 ; LENGTH: 267
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-09-917-254-101

Query Match 96.5%; Score 1344; DB 4; Length 267;
 Best Local Similarity 98.4%; Pred. No. 3.6e-136;
 Matches 241; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 64
 |||||
 DB 23 IVGGQEARPSKMPQVSLRVHGPYMMHFCGSSLIHPQWVLTAAACVGPVDVLDLAAALRVQL 82

QY 65 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 124
 |||||
 DB 83 REQHLYYODQLLPVSRITIVHPQFYTAQIGADIALLEBEPVKSSSHVTVTLPPASETFP 142

QY 125 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 184
 |||||
 DB 143 PGMPCWVTGMGDVNDRLPPFPPLKQVPI MENHI CDAKYHLGAYTGDDVRIVRDDML 202

QY 185 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 244
 |||||
 DB 203 CAGNTRRDSGCGDSSGGLVCKVNGTWMLOAGVSVSGEGCAQPNRPGIYTRVTVYLDWIHHY 262

QY 245 VPKKP 249
 |||||
 DB 263 VPKKP 267

RESULT 15
 US-08-944-483-69
 ; Sequence 69, Application US/08944483
 ; Patent No. 6232456
 ; GENERAL INFORMATION:
 ; APPLICANT: COHEN, MAURICE

Job time : 64 secs

APPLICANT: COLPITTS, TRACEY L.
 APPLICANT: FRIEDMAN, PAULA N.
 APPLICANT: GRANADOS, EDWARD N.
 APPLICANT: KASS, MICHAEL R.
 APPLICANT: RUSSELL, JOHN C.
 APPLICANT: STEWART, KENT D.
 APPLICANT: STROUPE, STEVEN D.
 TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
 TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
 TITLE OF INVENTION: OF THE PROSTATE
 NUMBER OF SEQUENCES: 76
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Abbott Laboratories
 STREET: 100 Abbott Park Road
 CITY: Abbott Park
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/944,483
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Becker, Cheryl L.
 REGISTRATION NUMBER: 35,441
 REFERENCE/DOCKET NUMBER: 6183.US.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847/935-1729
 TELEFAX: 847/938-2623
 TELEX:
 INFORMATION FOR SEQ ID NO: 69:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 245 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 6232456e
 US-08-944-483-69

Query Match 91.5%; Score 1274; DB 3; Length 245;
 Best Local Similarity 92.2%; Pred. No. 1.1e-128;
 Matches 226; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY	5	IVGGQEARPSKRWQVSLRVHNGPYMMHFCGGSIIHPQWVLTAAACVDPVKDIAALRVQL	64
DB	1	IVGGQEARPSKRWQVSLRVHNGPYMMHFCGGSIIHPQWVLTAAACVDPVKDIAALRVQL	60
QY	65	RQQLIYYDQLLPSRIIVHROPYTAQIGADIALLEBEPVKVSSHHTVTLPPASSTPP	124
DB	61	RQQLIYYDQLLPSRIIVHROPYTAQIGADIALLEBEPVKVSSHHTVTLPPASSTPP	120
QY	125	PGMPCWVTGMDVNDNERLPPRPRLKQVKVPIEMENHICDAKYNHLAGAYTGGDDVRIVRDML	184
DB	121	PGMPCWVTGMDVNDNERLPPRPRLKQVKVPIEMENHICDAKYNHLAGAYTGGDDVRIVRDML	180
QY	185	CAGNTRRDSQQDSDGSPVLCVKNVGTWLDAGVVSWSGCGAQPNRPGIYTRVTVYLLDWIHHY	244
DB	181	CAGNSQRDSCKGGSDGSPVLCVKNVGTWLDAGVVSWSGCGAQPNRPGIYTRVTVYLLDWIHHY	240
QY	245	VPKKP 249	
DB	241	VPKKP 245	

Search completed: August 27, 2005, 23:13:02

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 19:27:19 ; Search time 626 Seconds
(without alignments)
8058.662 Million cell updates/sec

Title: US-09-598-982C-20

Perfect score: 771
Sequence: 1 gggccccccgagaaagaat.....cgtgagcggccgcgcgcgt 771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 731713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*	PublSeq	PublLen	PublDesc
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*	14663426	200000000	US07_PUBCOMB
2: /cgn2_6/ptodata/2/pubpna/BCT_NEW_PUB.seq:*	14663426	200000000	BCT_NEW_PUB
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*	14663426	200000000	US06_NEW_PUB
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*	14663426	200000000	US06_PUBCOMB
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*	14663426	200000000	US07_NEW_PUB
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*	14663426	200000000	US07_PUBCOMB
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*	14663426	200000000	US08_NEW_PUB
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*	14663426	200000000	US08_PUBCOMB
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*	14663426	200000000	US09A_PUBCOMB
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*	14663426	200000000	US09B_PUBCOMB
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*	14663426	200000000	US09C_PUBCOMB
12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*	14663426	200000000	US09D_PUBCOMB
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*	14663426	200000000	US10A_PUBCOMB
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*	14663426	200000000	US10B_PUBCOMB
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*	14663426	200000000	US10C_PUBCOMB
16: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*	14663426	200000000	US10E_PUBCOMB
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*	14663426	200000000	US10F_PUBCOMB
18: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*	14663426	200000000	US10G_PUBCOMB
19: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*	14663426	200000000	US10H_PUBCOMB
20: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*	14663426	200000000	US10I_PUBCOMB
21: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*	14663426	200000000	US10J_PUBCOMB
22: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*	14663426	200000000	US11A_PUBCOMB
23: /cgn2_6/ptodata/2/pubpna/US11B_PUBCOMB.seq:*	14663426	200000000	US11B_PUBCOMB
24: /cgn2_6/ptodata/2/pubpna/US11C_PUBCOMB.seq:*	14663426	200000000	US11C_PUBCOMB
25: /cgn2_6/ptodata/2/pubpna/US11D_PUBCOMB.seq:*	14663426	200000000	US11D_PUBCOMB
26: /cgn2_6/ptodata/2/pubpna/US11E_PUBCOMB.seq:*	14663426	200000000	US11E_PUBCOMB

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	731	94.8	1143	US-10-352-684A-47	Sequence 47, Appl
2	731	94.8	1145	US-10-287-226-93	Sequence 93, Appl
3	731	94.8	1148	US-10-956-157-2444	Sequence 2444, Ap
4	731	94.8	1193	US-10-956-157-1043	Sequence 1043, Ap
5	730.4	94.7	828	US-10-287-226-91	Sequence 91, Appl
6	729.4	94.6	1334	US-10-723-860-6799	Sequence 6799, Ap
7	726.2	94.2	1081	US-09-954-456-2126	Sequence 2126, Ap

8	726.2	94.2	1081	10	US-09-960-706-680	Sequence 680, App
9	726.2	94.2	1081	10	US-09-873-319-427	Sequence 427, App
10	726.2	94.2	1081	10	US-09-873-367C-155	Sequence 155, App
11	726.2	94.2	1081	10	US-09-873-367C-714	Sequence 714, App
12	726.2	94.2	1081	21	US-10-843-641A-155	Sequence 155, App
13	726.2	94.2	1081	21	US-10-843-641A-714	Sequence 714, App
14	726.2	94.2	1081	21	US-10-843-641A-5153	Sequence 5153, Ap
15	718.2	93.2	2662	18	US-10-275-505-27	Sequence 27, Appl
16	700	90.8	828	18	US-10-287-226-89	Sequence 89, Appl
17	686.2	89.0	1154	14	US-10-352-684A-45	Sequence 45, Appl
18	686.2	89.0	1158	17	US-10-116-802-240	Sequence 240, App
19	686.2	89.0	1158	21	US-10-956-157-1042	Sequence 1042, Ap
20	678	87.9	858	18	US-10-287-226-87	Sequence 87, Appl
21	507.6	65.8	729	21	US-10-480-988-55	Sequence 55, Appl
22	498	64.6	729	21	US-10-956-157-1789	Sequence 1789, Ap
23	484.4	62.8	600	21	US-10-956-157-7024	Sequence 7024, Ap
24	463	60.1	1187	18	US-10-275-505-22	Sequence 22, Appl
25	380	49.3	5456	20	US-10-723-860-2689	Sequence 2689, Ap
26	380	49.3	5456	22	US-10-756-149-2484	Sequence 2484, Ap
27	287.6	37.3	600	21	US-10-956-157-6278	Sequence 6278, Ap
28	287.6	37.3	600	21	US-10-956-157-7679	Sequence 7679, Ap
29	287.4	37.3	846	10	US-09-813-432-11	Sequence 11, Appl
30	287.4	37.3	846	17	US-10-174-364-11	Sequence 11, Appl
31	287.4	37.3	846	18	US-10-246-583-11	Sequence 11, Appl
32	287.4	37.3	846	19	US-10-689-832-11	Sequence 11, Appl
33	275	35.7	600	10	US-09-918-995-6378	Sequence 6378, Ap
34	274.8	35.6	600	21	US-10-956-157-6277	Sequence 6277, Ap
35	252.4	32.7	948	16	US-10-131-409-15	Sequence 15, Appl
36	252.4	32.7	948	17	US-10-139-854-15	Sequence 15, Appl
37	252.4	32.7	948	17	US-10-150-813-15	Sequence 15, Appl
38	252.4	32.7	948	17	US-10-150-811-15	Sequence 15, Appl
39	236.4	30.7	858	16	US-10-131-409-21	Sequence 21, Appl
40	236.4	30.7	858	17	US-10-139-854-21	Sequence 21, Appl
41	236.4	30.7	858	17	US-10-150-813-21	Sequence 21, Appl
42	236.4	30.7	858	17	US-10-150-811-21	Sequence 21, Appl
43	233.4	30.3	873	14	US-10-117-323-2	Sequence 2, Appl1
44	231.6	30.0	843	16	US-10-131-409-25	Sequence 25, Appl1
45	231.6	30.0	843	17	US-10-139-854-25	Sequence 25, Appl1

ALIGNMENTS

RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Healy, Aileen
; APPLICANT: Welch, Nadine S.
; APPLICANT: Kelly, Louise M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
; FILE REFERENCE: MPI02-019PINNONMIM
; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
; CURRENT FILING DATE: US/10/352,684A
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/354,333
; PRIOR APPLICATION NUMBER: US 60/360,258
; PRIOR APPLICATION NUMBER: US 60/364,476
; PRIOR APPLICATION NUMBER: US 60/375,626
; PRIOR APPLICATION NUMBER: US 60/375,626
; PRIOR APPLICATION NUMBER: US 60/386,494
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/390,965
; PRIOR FILING DATE: 2002-06-24

```

? PRIOR APPLICATION NUMBER: US 60/392,480
? PRIOR FILING DATE: 2002-06-28
? PRIOR APPLICATION NUMBER: US 60/394,128
? PRIOR FILING DATE: 2002-07-03
? PRIOR APPLICATION NUMBER: US 60/399,783
? PRIOR FILING DATE: 2002-07-31
? PRIOR APPLICATION NUMBER: US 60/403,221
? PRIOR FILING DATE: 2002-08-13
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 62
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 47
? LENGTH: 1143
? TYPE: DNA
? ORGANISM: Homo Sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (6)...(833)
? US-10-352-684A-47

```

```

Query Match 94.8% Score 731; DB 17; Length 1143;
Best Local Similarity 98.0%; Pred. No. 2, 6e-195;
Matches 740; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 17 GAATTCGTCCGGGGGTCAAGAGAGCCCGCCAGAGCAAGTGGCCCTGGCAAGTGGAGAG 76
DB 94 GCATTCGTTGGGGGTCAAGAGAGCCCGCCAGAGCAAGTGGCCCTGGCAAGTGGAGAG 153
QY 77 TCACAGGGCCCATCTGATGATGATCTTGGCGGGGGCTCCCTCATCCACCCTCAAGTGG 136
DB 154 TCACAGGGCCCATCTGATGATGATCTTGGCGGGGGCTCCCTCATCCACCCTCAAGTGG 213
QY 137 TGAACCGCCGGGGGTGGGGGAGCCCGCCAGAGCAAGTGGCCCTGGCAAGTGGAGAG 196
DB 214 TGACCCGAGGACACTGGTGGAGACCGGAGCTCAAGATCTGGCCCTCAAGTGGAGAG 273
QY 197 TGGGGGAGAGCACTCTACTACAGAGCCAGTGTCTGGCGGATCAAGAGATCATGTGTG 256
DB 274 TGGGGGAGAGCACTCTACTACAGAGCCAGTGTCTGGCGGATCAAGAGATCATGTGTG 333
QY 257 ACCCAAGTCTACACCGCCAGATGGAGCGGAGCAATGGCCCTGGCAAGTGGAGAGAG 316
DB 334 ACCCAAGTCTACACCGCCAGATGGAGCGGAGCAATGGCCCTGGCAAGTGGAGAGAG 393
QY 317 CGGTGAAGTCTCCAGCCAGTCCACACCGTCCACCTGGCCCTGGCAAGTGGAGAG 376
DB 394 CGGTGAAGTCTCCAGCCAGTCCACACCGTCCACCTGGCCCTGGCAAGTGGAGAG 453
QY 377 CCCCAGGATGCGGTGTGGGTCACTGGTGGGGGATGTGACAAATGATGAGCGCTCC 436
DB 454 CCCCAGGATGCGGTGTGGGTCACTGGTGGGGGATGTGACAAATGATGAGCGCTCC 513
QY 437 CACCGCATTTCTCTGAAAGCAGTGAAGTCCCAATTAAGAAACCAATTTGAGAG 496
DB 514 CACCGCATTTCTCTGAAAGCAGTGAAGTCCCAATTAAGAAACCAATTTGAGAG 573
QY 497 CAAAATACACTTGGGCTTACACCGGAGAGCAGTCCGATCTCCGATGAGAGATGC 556
DB 574 CAAAATACACTTGGGCTTACACCGGAGAGCAGTCCGATCTCCGATGAGAGATGC 633
QY 557 TGTGTGCGGGGAAACACCCGGAGGAGACTCATGCAAGGGGCACTCCGAGAGGGCCCTGTGT 616
DB 634 TGTGTGCGGGGAAACACCCGGAGGAGACTCATGCAAGGGGCACTCCGAGAGGGCCCTGTGT 693
QY 617 GCNAAGTGAATGCACTTGGCTGTGAGGCGGGCGTGTCACTGGGGGCGAGGGGTGTGCC 676
DB 694 GCNAAGTGAATGCACTTGGCTGTGAGGCGGGCGTGTCACTGGGGGCGAGGGGTGTGCC 753
QY 677 AGCCCAACCGGCTGGATTAACCCCGTGTCACTACTACTTGAATGAGATCCACCCT 736
DB 754 AGCCCAACCGGCTGGATTAACCCCGTGTCACTACTACTTGAATGAGATCCACCCT 813
QY 737 ATGTCCCAAAAAGCCGTGAAGCGCCGCGTGTGT 771

```

```

DB 814 ATGTCCCAAAAAGCCGTGAAGCGCCGCGTGTGT 848

```

```

RESULT 2
US-10-287-226-93
? Sequence 93, Application US/10287226
? Publication No. US20040086875A1
? GENERAL INFORMATION:
? APPLICANT: Agee, Michele L.,
? APPLICANT: Alsbroock, John P.,
? APPLICANT: Bergns, Constance,
? APPLICANT: Boldog, Ference,
? APPLICANT: Burgess, Catherine E.,
? APPLICANT: Chant, John S.,
? APPLICANT: Chaudhuri, Amitabha,
? APPLICANT: DiPippo, Vincent A.,
? APPLICANT: Edinger, Salomit R.,
? APPLICANT: Eilsen, Andrew,
? APPLICANT: Ellerman, Karen,
? APPLICANT: Gangolli, Esha A.,
? APPLICANT: Gorman, Linda,
? APPLICANT: Gerlach, Valerie,
? APPLICANT: Ji, Weizhen,
? APPLICANT: Kekuda, Ramesh,
? APPLICANT: Khrantsev, Nikolai,
? APPLICANT: Li, Li,
? APPLICANT: Maljankar, Uriel M.,
? APPLICANT: Machougall, John R.,
? APPLICANT: Mezes, Peter S.,
? APPLICANT: Miller, Charles E.,
? APPLICANT: Millet, Isabelle,
? APPLICANT: Ooi, Chean Eng,
? APPLICANT: Ort, Tatiana,
? APPLICANT: Padigaru, Muralihara,
? APPLICANT: Paturajan, Meera,
? APPLICANT: Rastelli, Luca,
? APPLICANT: Rieger, Daniel K.,
? APPLICANT: Rothenberg, Mark E.,
? APPLICANT: Shenoy, Suresh G.,
? APPLICANT: Spaderna, Steven K.,
? APPLICANT: Spylek, Kimberley A.,
? APPLICANT: Taupier, Jr., Raymond J.,
? APPLICANT: Vernet, Corine A.M.,
? APPLICANT: Zernusen, Bryan D.,
? APPLICANT: Zhong, Mei
? TITLE OR INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
? FILE REFERENCE: 21402-480C
? CURRENT FILING DATE: 2002-11-04
? PRIOR APPLICATION NUMBER: 60/334,421
? PRIOR FILING DATE: 2001-11-30
? PRIOR APPLICATION NUMBER: 60/354,392
? PRIOR FILING DATE: 2002-02-04
? PRIOR APPLICATION NUMBER: 60/360,148
? PRIOR FILING DATE: 2002-02-27
? PRIOR APPLICATION NUMBER: 60/364,000
? PRIOR FILING DATE: 2002-03-13
? PRIOR APPLICATION NUMBER: 60/404,821
? PRIOR FILING DATE: 2002-08-20
? PRIOR APPLICATION NUMBER: 60/334,526
? PRIOR FILING DATE: 2001-11-30
? PRIOR APPLICATION NUMBER: 60/354,409
? PRIOR FILING DATE: 2002-02-04
? PRIOR APPLICATION NUMBER: 60/364,227
? PRIOR FILING DATE: 2002-03-13
? PRIOR APPLICATION NUMBER: 60/334,027
? PRIOR FILING DATE: 2001-11-28
? PRIOR APPLICATION NUMBER: 60/331,641
? PRIOR FILING DATE: 2001-11-20
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 673
? SOFTWARE: Cursesqlist version 0.1

```

```

; SEQ ID NO 93
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)..(832)
US-10-287-226-93

```

```

Query Match 94.8%; Score 731; DB 18; Length 1145;
Best Local Similarity 98.7%; Pred. No. 2.6e-195;
Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 17 GAATGTCGGGGGTCAGAGAGGCCCCCAGAGAGCAATGGCCCTGGAGGTGAGCCCTGAGAG 76
DB 96 GCATCGTTGGGGGTCAAGAGAGGCCCCCAGAGAGCAATGGCCCTGGAGGTGAGCCCTGAGAG 155
OY 77 TCCAGCGCCCACTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 136
DB 156 TCCAGCGCCCACTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 215
OY 137 TGACCGCGCCGCGCGTGGGAGCCGGAAGTCAAGGATCTGGCCGCTCAAGGTGCAAC 196
DB 216 TGACCGCGCCGCGCGTGGGAGCCGGAAGTCAAGGATCTGGCCGCTCAAGGTGCAAC 275
OY 197 TGCCGGAGCAGACCTCTACTACAGAGACCAAGCTGCTGCGGTCAAGAGATCACTGTC 256
DB 276 TGCCGGAGCAGACCTCTACTACAGAGACCAAGCTGCTGCGGTCAAGAGATCACTGTC 335
OY 257 ACCCAGAGTTCTACACCGCCAGATCGAGAGGAGCATGCGCCCTGCTGAGGTGAGAGC 316
DB 336 ACCCAGAGTTCTACACCGCCAGATCGAGAGGAGCATGCGCCCTGCTGAGGTGAGAGC 395
OY 317 CCGTGAAGGTCTCCAGCCAGTCCAACGATCAACGATCAACGATCAACGATCAACGATCA 376
DB 396 CCGTGAAGGTCTCCAGCCAGTCCAACGATCAACGATCAACGATCAACGATCAACGATCA 455
OY 377 CCGGGGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB 456 CCGGGGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 515
OY 437 CACCGCAATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 496
DB 516 CACCGCAATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 575
OY 497 CAAAATACCACTTGGGCGCTTACACGGGAGCAAGCTCCGATCGTCCGTGACGATGCG 556
DB 576 CAAAATACCACTTGGGCGCTTACACGGGAGCAAGCTCCGATCGTCCGTGACGATGCG 635
OY 557 TGTGTGCGGGGAAKACCGGGAGGACTCATGCGAGGGAGCTCCGGAGGGCCCTGTGTGT 616
DB 636 TGTGTGCGGGGAAKACCGGGAGGACTCATGCGAGGGAGCTCCGGAGGGCCCTGTGTGT 695
OY 617 GCAAGGTGAATGCACTGAGTGCAGGCGGGGCTGATGCTGAGGAGGAGGCTGTGCCC 676
DB 696 GCAAGGTGAATGCACTGAGTGCAGGCGGGGCTGATGCTGAGGAGGAGGCTGTGCCC 755
OY 677 AGCCCAACCGGCTGGCATCTACCCCGTGCACCTACTACTTGGATCGATCGACACT 736
DB 756 AGCCCAACCGGCTGGCATCTACCCCGTGCACCTACTACTTGGATCGATCGACACT 815
OY 737 ATGTCCCAAAAAGCCGTGAAGCGGC 763
DB 816 ATGTCCCAAAAAGCCGTGAAGCGGC 842

```

```

RESULT 3
US-10-956-157-2444
; Sequence 2444, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William

```

```

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2444
; LENGTH: 1148
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2444

```

```

Query Match 94.8%; Score 731; DB 21; Length 1148;
Best Local Similarity 98.7%; Pred. No. 2.6e-195;
Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 17 GAATGTCGGGGGTCAGAGAGGCCCCCAGAGAGCAATGGCCCTGGAGGTGAGCCCTGAGAG 76
DB 96 GCATCGTTGGGGGTCAAGAGAGGCCCCCAGAGAGCAATGGCCCTGGAGGTGAGCCCTGAGAG 155
OY 77 TCCAGCGCCCACTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 136
DB 156 TCCAGCGCCCACTACTGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 215
OY 137 TGACCGCGCCGCGCGTGGGAGCCGGAAGTCAAGGATCTGGCCGCTCAAGGTGCAAC 196
DB 216 TGACCGCGCCGCGCGTGGGAGCCGGAAGTCAAGGATCTGGCCGCTCAAGGTGCAAC 275
OY 197 TGCCGGAGCAGACCTCTACTACAGAGACCAAGCTGCTGCGGTCAAGAGATCACTGTC 256
DB 276 TGCCGGAGCAGACCTCTACTACAGAGACCAAGCTGCTGCGGTCAAGAGATCACTGTC 335
OY 257 ACCCAGAGTTCTACACCGCCAGATCGAGAGGAGCATGCGCCCTGCTGAGGTGAGAGC 316
DB 336 ACCCAGAGTTCTACACCGCCAGATCGAGAGGAGCATGCGCCCTGCTGAGGTGAGAGC 395
OY 317 CCGTGAAGGTCTCCAGCCAGTCCAACGATCAACGATCAACGATCAACGATCAACGATCA 376
DB 396 CCGTGAAGGTCTCCAGCCAGTCCAACGATCAACGATCAACGATCAACGATCAACGATCA 455
OY 377 CCGGGGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB 456 CCGGGGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 515
OY 437 CACCGCAATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 496
DB 516 CACCGCAATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 575
OY 497 CAAAATACCACTTGGGCGCTTACACGGGAGCAAGCTCCGATCGTCCGTGACGATGCG 556
DB 576 CAAAATACCACTTGGGCGCTTACACGGGAGCAAGCTCCGATCGTCCGTGACGATGCG 635
OY 557 TGTGTGCGGGGAAKACCGGGAGGACTCATGCGAGGGAGCTCCGGAGGGCCCTGTGTGT 616
DB 636 TGTGTGCGGGGAAKACCGGGAGGACTCATGCGAGGGAGCTCCGGAGGGCCCTGTGTGT 695
OY 617 GCAAGGTGAATGCACTGAGTGCAGGCGGGGCTGATGCTGAGGAGGAGGCTGTGCCC 676
DB 696 GCAAGGTGAATGCACTGAGTGCAGGCGGGGCTGATGCTGAGGAGGAGGCTGTGCCC 755
OY 677 AGCCCAACCGGCTGGCATCTACCCCGTGCACCTACTACTTGGATCGATCGACACT 736
DB 756 AGCCCAACCGGCTGGCATCTACCCCGTGCACCTACTACTTGGATCGATCGACACT 815
OY 737 ATGTCCCAAAAAGCCGTGAAGCGGC 763
DB 816 ATGTCCCAAAAAGCCGTGAAGCGGC 842

```

```

RESULT 4
US-10-956-157-1043
; Sequence 1043, Application US/10956157

```

```

; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeath
; APPLICANT: Mounst, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1043
; LENGTH: 1193
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-956-157-1043

```

```

Query Match 94.8% Score 731, DB 21, Length 1193;
Best Local Similarity 98.7%; Pred. No. 2,6e-195;
Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy 17 GAATCGCGGGGTCAGAGAGGCCCCAGAGCAAGTGAGCCCTGCGCAGGATGAGCCGTGAGAG 75
Db 141 GCATCGTCGGGGGTGAGAGAGCCCCAGAGCAAGTGAGCCCTGCGCAGGATGAGCCGTGAGAG 200
Qy 77 TCCAAGGCCGATATCTGATGATCTGATGATCTGATGATCTGATGATCTGATGATCTGATG 136
Db 201 TCCAAGGCCGATATCTGATGATCTGATGATCTGATGATCTGATGATCTGATGATCTGATG 250
Qy 137 TGAACCGCCGGGGGTGATGAGGAGCGGAGCCGATGAGGATCTGATGATGAGCGCCGTGAGAG 196
Db 261 TGAACCGGAGGAGCGGATGATGAGGAGCGGAGCCGATGAGGATCTGATGATGAGCGCCGTGAGAG 320
Qy 197 TGGGGGAGCAGCAGCTCTATACAGAGGAGGATGATGATGATGATGATGATGATGATGATG 256
Db 321 TGGGGGAGCAGCAGCTCTATACAGAGGAGGATGATGATGATGATGATGATGATGATGATG 380
Qy 257 ACCCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 316
Db 381 ACCCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 440
Qy 317 CGGMAAGGCTCCAGCAGCAGCTCCAGCAGCAGCTCCAGCAGCAGCTCCAGCAGCAGCTCC 376
Db 441 CGGMAAGGCTCCAGCAGCAGCTCCAGCAGCAGCTCCAGCAGCAGCTCCAGCAGCAGCTCC 500
Qy 377 CCCCAGGAGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 436
Db 501 CCCCAGGAGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
Qy 437 CACCGCATTTCTCTGAGCAGGATGATGATGATGATGATGATGATGATGATGATGATG 496
Db 561 CACCGCATTTCTCTGAGCAGGATGATGATGATGATGATGATGATGATGATGATGATG 620
Qy 497 GAAAATPACACCTTGGGGCCCTAGACGGGAGAGCAACGCGCCGATGATGATGATGATGATG 556
Db 621 GAAAATPACACCTTGGGGCCCTAGACGGGAGAGCAACGCGCCGATGATGATGATGATGATG 680
Qy 557 TGTGTGCCCCGGAACAACCCGGAGGAGTCCATGCCAGGCGGATCCCGAGGGCCCTGATG 616
Db 681 TGTGTGCCCCGGAACAACCCGGAGGAGTCCATGCCAGGCGGATCCCGAGGGCCCTGATG 740
Qy 617 GCAAAGTGAATGGATCCTGCTGATGAGGCGGCGGATGATGATGATGATGATGATGATG 676
Db 741 GCAAAGTGAATGGATCCTGCTGATGAGGCGGCGGATGATGATGATGATGATGATGATG 800
Qy 677 AGCCCAACGGGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 736
Db 801 AGCCCAACGGGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860
Qy 737 ATGTCCCAGAAAAAGCGGTGAGAGCGCC 763
Db 861 ATGTCCCAGAAAAAGCGGTGAGAGCGC 887

```

```

RESULT 5
US-10-287-226-91
; Sequence 91, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsbrock, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomo R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Elleman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khramtsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyanek, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patlurajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suren G.,
; APPLICANT: Spaderma, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Wei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: Cursesqlist version 0.1
; SEQ ID NO 91
; LENGTH: 828
; TYPE: DNA

```


ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(825)
US-10-287-226-91

Query Match 94.7%: Score 730.4; DB 18; Length 828;
Best Local Similarity 99.2%: Pred. No. 3.7e-195;
Matches 734; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 17 GAATCGTGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
DB GGATCGTGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 148
OY 77 TCCAGGCCCCAATCTGATGCACTTCTGGGGGGGCTCCCTCATCCACCCTCCAGTGGTGC 136
DB TCCAGGCCCCAATCTGATGCACTTCTGGGGGGGCTCCCTCATCCACCCTCCAGTGGTGC 208
OY 137 TGACCCGGCCGGCGTGGGGGAGCCGGACGTCAAGGANTCTGGCCGCTCAGGGTCAAC 196
DB TGACCCGGCCGGCGTGGGGGAGCCGGACGTCAAGGANTCTGGCCGCTCAGGGTCAAC 268
OY 197 TCGGGAGCAGCACTCTACTACAGAGCAGCTGCTGCGGCTCAGCAGATCATCTGTC 256
DB TCGGGAGCAGCACTCTACTACAGAGCAGCTGCTGCGGCTCAGCAGATCATCTGTC 328
OY 257 ACCCAAGTTTAAACCCGCCAGATCGGAGCGGACATGCGCCCTGCTGAGCTGAGAGC 316
DB ACCCAAGTTTAAACCCGCCAGATCGGAGCGGACATGCGCCCTGCTGAGCTGAGAGC 388
OY 317 CGGTGAAGTCTCCAGCCAGGTCACAGGTGACCCCTGCGCTCCAGAGACCTTCC 376
DB CGGTGAAGTCTCCAGCCAGGTCACAGGTGACCCCTGCGCTCCAGAGACCTTCC 448
OY 377 CCCCCGGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
DB CCCCCGGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 508
OY 437 CACGGCCATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 496
DB CACGGCCATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 568
OY 497 CAAAATPACCACTTGGCCGCTCAACAGGAGAGAGAGTCCCGCANTGCTGAGAGCATG 556
DB CAAAATPACCACTTGGCCGCTCAACAGGAGAGAGAGTCCCGCANTGCTGAGAGCATG 628
OY 557 TGTGTGCCGGGAAACCCGGAGGACTCATGCGCAGGGGCTCCGGAGGGCCCTGCTGT 616
DB TGTGTGCCGGGAAACCCGGAGGACTCATGCGCAGGGGCTCCGGAGGGCCCTGCTGT 688
OY 617 GCAAGGTGAATGGCACTGCTGCAAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
DB GCAAGGTGAATGGCACTGCTGCAAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
OY 677 AGCCCAACCGGCTGGGATCTACACCCGGTCACTCACTTGAATGAGATCCACCACT 736
DB AGCCCAACCGGCTGGGATCTACACCCGGTCACTCACTTGAATGAGATCCACCACT 808
OY 737 ATGTCCCAAAAAGCCGTGA 756
DB ATGTCCCAAAAAGCCGTGA 828

RESULT 6
US-10-723-860-6799
Sequence 6799, Application US/10723860
Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Ginsburg, Wendy M.
APPLICANT: Zlotnick, Albert
TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators

FILE REFERENCE: 05882.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723.860
CURRENT FILING DATE: 2003-11-26
PRIOR APPLICATION NUMBER: 60/429.739
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 6799
LENGTH: 1334
TYPE: DNA
ORGANISM: Homo sapiens
US-10-723-860-6799

Query Match 94.6%: Score 729.4; DB 20; Length 1334;
Best Local Similarity 98.5%: Pred. No. 7.5e-195;
Matches 736; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 17 GAATCGTGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
DB GGATCGTGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 200
OY 77 TCCAGGCCCCAATCTGATGCACTTCTGGGGGGGCTCCCTCATCCACCCTCCAGTGGTGC 136
DB TCCAGGCCCCAATCTGATGCACTTCTGGGGGGGCTCCCTCATCCACCCTCCAGTGGTGC 260
OY 137 TGACCCGGCCGGCGTGGGGGAGCCGGACGTCAAGGANTCTGGCCGCTCAGGGTCAAC 196
DB TGACCCGGCCGGCGTGGGGGAGCCGGACGTCAAGGANTCTGGCCGCTCAGGGTCAAC 320
OY 197 TCGGGAGCAGCACTCTACTACAGAGCAGCTGCTGCGGCTCAGCAGATCATCTGTC 256
DB TCGGGAGCAGCACTCTACTACAGAGCAGCTGCTGCGGCTCAGCAGATCATCTGTC 380
OY 257 ACCCAAGTTTAAACCCGCCAGATCGGAGCGGACATGCGCCCTGCTGAGCTGAGAGC 316
DB ACCCAAGTTTAAACCCGCCAGATCGGAGCGGACATGCGCCCTGCTGAGCTGAGAGC 440
OY 317 CGGTGAAGTCTCCAGCCAGGTCACAGGTGACCCCTGCGCTCCAGAGACCTTCC 376
DB CGGTGAAGTCTCCAGCCAGGTCACAGGTGACCCCTGCGCTCCAGAGACCTTCC 500
OY 377 CCCCCGGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
DB CCCCCGGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
OY 437 CACGGCCATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 496
DB CACGGCCATTTCTCTGAAAGCAGGTGAAAGTCCCAATATGAAACCAATTTGTGACG 620
OY 497 CAAAATPACCACTTGGCCGCTCAACAGGAGAGAGAGTCCCGCANTGCTGAGAGCATG 556
DB CAAAATPACCACTTGGCCGCTCAACAGGAGAGAGAGTCCCGCANTGCTGAGAGCATG 680
OY 557 TGTGTGCCGGGAAACCCGGAGGACTCATGCGCAGGGGCTCCGGAGGGCCCTGCTGT 616
DB TGTGTGCCGGGAAACCCGGAGGACTCATGCGCAGGGGCTCCGGAGGGCCCTGCTGT 740
OY 617 GCAAGGTGAATGGCACTGCTGCAAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
DB GCAAGGTGAATGGCACTGCTGCAAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
OY 677 AGCCCAACCGGCTGGGATCTACACCCGGTCACTCACTTGAATGAGATCCACCACT 736
DB AGCCCAACCGGCTGGGATCTACACCCGGTCACTCACTTGAATGAGATCCACCACT 860
OY 737 ATGTCCCAAAAAGCCGTGAAGCCGTC 763
DB ATGTCCCAAAAAGCCGTGAAGCCGTC 887

RESULT 7
US-09-954-456-2126
Sequence 2126, Application US/09954456

```

; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-2126

Query Match      94.2%; Score 726.2; DB 9; Length 1081;
Best Local Similarity 98.3%; Pred. No. 5,8e-194;
Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasi
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
; US-09-960-706-680

Query Match      94.2%; Score 726.2; DB 10; Length 1081;
Best Local Similarity 98.3%; Pred. No. 5,8e-194;
Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

Oy 377 CCCCCGGGGGATGCCGTCGTGGGTCACTGATGCGGGAGATGTGGAACAATGAGAGCCCTCC 436
Db 426 CCCCCGGGGGATGCCGTCGTGGGTCACTGATGCGGGAGATGTGGAACAATGAGAGCCCTCC 485
Oy 437 CACCGCCATTTTCTCTCTGAAAGCAGTGAAGGTCCCACTAATGAAAAACAATTTGTCAGC 496
Db 486 CACCGCCATTTTCTCTCTGAAAGCAGTGAAGGTCCCACTAATGAAAAACAATTTGTCAGC 545
Oy 497 CAAAATACCACTTGTGGCGCCCTCAACAGGGGAAGAGAGTCCCGCATTCCTGTGCAAGCATGC 556
Db 546 CAAAATACCACTTGTGGCGCCCTCAACAGGGGAAGAGAGTCCCGCATTCCTGTGCAAGCATGC 605
Oy 557 TGTGTGCGGGGAACAACCGGGAAAGTCAATGCAAGGGGCACTCCGGAGAGGGCCCTGTGAT 616
Db 606 TGTGTGCGGGGAACAACCGGGAAAGTCAATGCAAGGGGCACTCCGGAGAGGGCCCTGTGAT 665
Oy 617 GCAAGGTGAATGGCAGTGTGGTCAGAGCGGGCGTGTGTCAGTCAGTCGAGGGGCTGTGCC 676
Db 666 GCAAGGTGAATGGCAGTGTGGTCAGAGCGGGCGTGTGTCAGTCAGTCGAGGGGCTGTGCC 725
Oy 677 AGCCCAACCGGCTGGCATCTACACCCCGTGTCACTACTTGAATCTGAGATCCAGCAT 736
Db 726 AGCCCAACCGGCTGGCATCTACACCCCGTGTCACTACTTGAATCTGAGATCCAGCAT 785
Oy 737 ATGTCCCCAATAAACCGGTAAGCGGCC 763
Db 786 ATGTCCCCAATAAACCGGTAAGCGGCC 812

```

RESULT 9

```

US-09-873-319-427
; Sequence 427, Application US/0987319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 427
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427

```

```

Query Match 94.2%; Score 726.2; DB 10; Length 1081;
Best Local Similarity 98.3%; Pred. No. 5,8e-194;
Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Oy 17 GAAATGTCGGGGGTGAGGAGGCGCCAGGAAGAAGTGGCCCGGAGGAGGAGCCGTCAGAG 76
Db 66 GAAATGTCGGGGGTGAGGAGGCGCCAGGAAGAAGTGGCCCGGAGGAGGAGCCGTCAGAG 125
Oy 77 TCACAGCCATATGATGATGACTTCTGCGGGGGTCCCTCAACACCCCAAGGATGTC 136
Db 126 TCACAGCCATATGATGATGACTTCTGCGGGGGTCCCTCAACACCCCAAGGATGTC 185
Oy 137 TCACCGCGGGCGGCGGTCGTCGGAACCGGACGTCAGGATGTGCGCGGCTGACGGTCGAC 196
Db 186 TCACCGCGGGCGGCGGTCGTCGGAACCGGACGTCAGGATGTGCGCGGCTGACGGTCGAC 245
Oy 197 TCGGGAGAGAGGAGCACTTAATCAAGGAGCAAGGATGTCGGGTGACGAGATCATTCGTGC 256

```

```

Db 246 TCGGGAGAGAGGAGCACTTAATCAAGGAGCAAGGATGTCGGGTGACGAGATCATTCGTGC 305
Oy 257 ACCCAAGTTCATCACCGCCCAAGTCGAGCGGACATTCGCGTCGTCGAGTCGAGAGC 316
Db 306 ACCCAAGTTCATCACCGCCCAAGTCGAGCGGACATTCGCGTCGTCGAGTCGAGAGC 365
Oy 317 CGGTGAAGTTCATCACCGTCAGTCCACAGCGTCAACCCGCTCCGCTCCAGAGACCTTCC 376
Db 366 CGGTGAAGTTCATCACCGTCAGTCCACAGCGTCAACCCGCTCCGCTCCAGAGACCTTCC 425
Oy 377 CCCCCGGGGATGCCGTCGTGGGTCACTGATGCGGGAGATGTGGAACAATGAGAGCCCTCC 436
Db 426 CCCCCGGGGATGCCGTCGTGGGTCACTGATGCGGGAGATGTGGAACAATGAGAGCCCTCC 485
Oy 437 CACCGCCATTTTCTCTGAAAGCAGTGAAGGTCCCACTAATGAAAAACAATTTGTCAGC 496
Db 486 CACCGCCATTTTCTCTGAAAGCAGTGAAGGTCCCACTAATGAAAAACAATTTGTCAGC 545
Oy 497 CAAAATACCACTTGTGGCGCCCTCAACAGGGGAAGAGAGTCCCGCATTCCTGTGCAAGCATGC 556
Db 546 CAAAATACCACTTGTGGCGCCCTCAACAGGGGAAGAGAGTCCCGCATTCCTGTGCAAGCATGC 605
Oy 557 TGTGTGCGGGGAACAACCGGGAAAGTCAATGCAAGGGGCACTCCGGAGAGGGCCCTGTGAT 616
Db 606 TGTGTGCGGGGAACAACCGGGAAAGTCAATGCAAGGGGCACTCCGGAGAGGGCCCTGTGAT 665
Oy 617 GCAAGGTGAATGGCAGTGTGGTCAGAGCGGGCGTGTGTCAGTCAGTCGAGGGGCTGTGCC 676
Db 666 GCAAGGTGAATGGCAGTGTGGTCAGAGCGGGCGTGTGTCAGTCAGTCGAGGGGCTVTGCC 725
Oy 677 AGCCCAACCGGCTGGCATCTACACCCCGTGTCACTACTTGAATCTGAGATCCAGCAT 736
Db 726 AGCCCAACCGGCTGGCATCTACACCCCGTGTCACTACTTGAATCTGAGATCCAGCAT 785
Oy 737 ATGTCCCCAATAAACCGGTAAGCGGCC 763
Db 786 ATGTCCCCAATAAACCGGTAAGCGGCC 812

```

RESULT 10

```

US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; EARLIER FILING DATE: 2003-04-29
; EARLIER APPLICATION NUMBER: U.S. 60/236,891
; NUMBER OF SEQ ID NOS: 2000-09-29
; NUMBER OF SEQ ID NOS: 60/236,842
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 155
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-155

```

```

Query Match 94.2%; Score 726.2; DB 10; Length 1081;

```

Best Local Similarity 98.3%; Pred. No. 5.8e-194; Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Table of sequence alignments for Query (QY) and Database (DB) entries, showing similarity scores and sequence identifiers.

US-09-873-367C-714; Sequence 714, Application US/09873367C; Publication No. US20030165839A1; GENERAL INFORMATION:

APPLICANT: Young, Paul; APPLICANT: Soppet, Daniel; APPLICANT: Andrews, Gregory; APPLICANT: Augustus, Meena; APPLICANT: Ebner, Reinhard; APPLICANT: Carter, Kenneth; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using...

PRIOR FILING DATE: 2000-09-29; PRIOR APPLICATION NUMBER: U.S. 60/236,842; PRIOR FILING DATE: 2000-09-29; PRIOR APPLICATION NUMBER: U.S. 60/244,867; PRIOR FILING DATE: 2000-11-01; PRIOR APPLICATION NUMBER: U.S. 60/245,084; NUMBER OF SEQ ID NOS: 1067; SOFTWARE: PatentIn version 3.0; SEQ ID NO 714; LENGTH: 1081; TYPE: DNA; ORGANISM: Homo sapiens; US-09-873-367C-714

Query Match 94.2%; Score 726.2; DB 10; Length 1081; Best Local Similarity 98.3%; Pred. No. 5.8e-194; Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Table of sequence alignments for Query (QY) and Database (DB) entries, showing similarity scores and sequence identifiers.

APPLICANT: Soppet, Daniel; APPLICANT: Andrews, Gregory; APPLICANT: Augustus, Meena; APPLICANT: Ebner, Reinhard; APPLICANT: Carter, Kenneth; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using...; FILE REFERENCE: 689290-64; CURRENT APPLICATION NUMBER: US/09/873,367C; CURRENT FILING DATE: 2003-04-29; PRIOR APPLICATION NUMBER: U.S. 60/236,891

```
US-10-843-641A-155
; Sequence 155, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-155
```

```
Query Match          94.2%; Score 726.2; DB 21; Length 1081;
Best Local Similarity 98.3%; Pred. No. 5.8e-194;
Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTCAAGGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGGCTGACCGTGAAG 76
DB 66 GCATCGTGGGGGTCAAGGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGGCTGACCGTGAAG 125
QY 77 TCCACGGCCCATTAAGTGAAGCACTTCTGGGGGGGCTCCCTCAATCCAAGGCTGAGGCTGC 136
DB 126 TCCGCACCGAATTAAGTGAAGCACTTCTGGGGGGGCTCCCTCAATCCAAGGCTGAGGCTGC 185
QY 137 TGAACCCCGCGGGGTGCGTGGAGCCGAGCAAGGATTTGGCCCGCTCAAGGGTGAAC 196
DB 186 TAGCCGCAAGCGCACTGCGTGGAGCCGAGCAAGGATTTGGCCCGCTCAAGGGTGAAC 245
QY 197 TGGGGAGGACGACCTTCTTAACAGGACCAAGTGTGTCGGCCGGTCAAGAGATTCATCTG 256
DB 246 TGGGGAGGACGACCTTCTTAACAGGACCAAGTGTGTCGGCCGGTCAAGAGATTCATCTG 305
QY 257 ACCCAAGTTCTAACAAGCGGCAAGTGGAGCGGCAATGGCCCTGGAGGCTGAGAGGAGC 316
DB 306 ACCCAAGTTCTAACAAGCGGCAAGTGGAGCGGCAATGGCCCTGGAGGCTGAGAGGAGC 365
QY 317 CGGTGAAGTCTTCAAGCCAGTCCACACGATCAAGGCTGCGCCCTGACGAGACCTTCC 376
DB 366 CGGTGAAGTCTTCAAGCCAGTCCACACGATCAAGGCTGCGCCCTGACGAGACCTTCC 425
QY 377 CCGCGGAGATCGCGTGGGTGATCGTGGGGGGGATGTGACATATATGAGCGCTTCC 436
DB 426 CCGCGGAGATCGCGTGGGTGATCGTGGGGGGGATGTGACATATATGAGCGCTTCC 485
QY 437 CACCGCATTTCTTGAAGCAGGTAAGGTCCCAATAATGAAAAACAATTTGTGACG 496
```

```
DB 486 CACCGCATTTCTTGAAGCAGGTAAGGTCCCAATAATGAAAAACAATTTGTGACG 545
QY 497 CAATAATCAACTTTGGGGCTTACACCGGAGACGAGCGTCCGATGCTCGTGAACAGATGC 556
DB 546 CAATAATCAACTTTGGGGCTTACACCGGAGACGAGCGTCCGATGCTCGTGAACAGATGC 605
QY 557 TGTGTCGGGGAAACACCCGAGGAGGACTCATGCGGAGGAGCCCTCGGTGTG 616
DB 606 TGTGTCGGGGAAACACCCGAGGAGGACTCATGCGGAGGAGCCCTCGGTGTG 665
QY 617 GCAAGGTGAATGACCTTGCGTGAAGCGGCGGTGTGACTGCTGGAGGGGAGGAGGCTGTCC 676
DB 666 GCAAGGTGAATGACCTTGCGTGAAGCGGCGGTGTGACTGCTGGAGGGGAGGAGGCTGTCC 725
QY 677 AGCCCAACCGGCTGGATTAACAACCGGTGATCACTTGGAGCTGGATCCAGCACT 736
DB 726 AGCCCAACCGGCTGGATTAACAACCGGTGATCACTTGGAGCTGGATCCAGCACT 785
QY 737 ATGTCCCAAAAAGCCGTGAAGCGGCC 763
DB 786 ATGTCCCAAAAAGCCGTGAAGTCAAGC 812
```

```
RESULT 13
US-10-843-641A-714
; Sequence 714, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-714

Query Match          94.2%; Score 726.2; DB 21; Length 1081;
Best Local Similarity 98.3%; Pred. No. 5.8e-194;
Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTCAAGGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGGCTGACCGTGAAG 76
DB 66 GCATCGTGGGGGTCAAGGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGGCTGACCGTGAAG 125
QY 77 TCCACGGCCCATTAAGTGAAGCACTTCTGGGGGGGCTCCCTCAATCCAAGGCTGAGGCTGC 136
```

Db 126 TCCGCCACCGATACTGATGACCTTTCGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 185
 QY 137 TGAACCGCCGCGGCGTGGTGGGAGACCGGACGTCGAAGATCTGGCCCGCCCTCAGGGTCAAC 196
 Db 186 TGAACCGGAGCGCACTGGTGGGAGACCGGACGTCGAAGATCTGGCCCGCCCTCAGGGTCAAC 245
 QY 197 TGGGGAGGACGACCTCTACTACAGGACCAAGCTGCTGGCGGTCAGAGGATCATGTGTGC 256
 Db 246 TCGGGGAGGACGACCTCTACTACAGGACCAAGCTGCTGGCGGTCAGAGGATCATGTGTGC 305
 QY 257 ACCCACAGTCTACACCCGCCAGATCGGAGCGGACATCGCCCTGCTGAGAGCTGAGAGG 316
 Db 306 ACCCACAGTCTCTACACCCGCCAGATCGGAGCGGACATCGCCCTGCTGAGAGCTGAGAGG 365
 QY 317 CGGTGAAGTCTCTCAAGCAGTCACACGCTGACCCCTGCCCTCCAGAGACCTTCC 376
 Db 366 CGGTGAAGTCTCTCAAGCAGTCACACGCTGACCCCTGCCCTCCAGAGACCTTCC 425
 QY 377 CCCCAGGAGATCGCTGCTGGGTCACTGGCTGGGGCGATGTGACATGATGAGCCCTCC 436
 Db 426 CCCCAGGAGATCGCTGCTGGGTCACTGGCTGGGGCGATGTGACATGATGAGCCCTCC 485
 QY 437 CACCGCATTTCTCTGAAACAGAGTGAAGGTCCCAATGAAACCAACATTTGTGACG 496
 Db 486 CACCGCATTTCTCTGAAACAGAGTGAAGGTCCCAATGAAACCAACATTTGTGACG 545
 QY 497 CAAAAATACACTTTGGCGCTTACACGGAGACGACGTCGCGATGTCCTGTGACGACATGC 556
 Db 546 CAAAAATACACTTTGGCGCTTACACGGAGACGACGTCGCGATGTCCTGTGACGACATGC 605
 QY 557 TGTGTGCGGGGAAACACCCCGGAGGAGCTCATGCGAGGGCGAATCCCGAAGGGCCCTGTGT 616
 Db 606 TGTGTGCGGGGAAACACCCCGGAGGAGCTCATGCGAGGGCGAATCCCGAAGGGCCCTGTGT 665
 QY 617 GCAAGGTGAATGAGCAGCTGCGTCAAGCGGGCGTGGTCACTGAGGGCGAGAGGCTGTGCC 676
 Db 666 GCAAGGTGAATGAGCAGCTGCGTCAAGCGGGCGTGGTCACTGAGGGCGAGAGGCTGTGCC 725
 QY 677 AGCCCAACCGGCGCTGGGATTAACACCGCGTGTACTTACTTGAATGAGCTGACCAACT 736
 Db 726 AGCCCAACCGGCGCTGGGATTAACACCGCGTGTACTTACTTGAATGAGCTGACCAACT 785
 QY 737 ATGTCCCAAAAAAGCCGTGAAGCGGCC 763
 Db 786 ATGTCCCAAAAAAGCCGTGAAGCGGCC 812

RESULT 14
 ; Sequence 5153, Application US/10843641A
 ; Publication No. US2005006454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ayalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE REFERENCE: 689290-189
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; CURRENT FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US/09/967,768
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/968,007

; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,347
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,708
 ; PRIOR FILING DATE: 2001-10-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 8447
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5153
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-843-641A-5153

Query Match 94.2%; Score 726.2; DB 21; Length 1081;
 Best Local Similarity 98.3%; Pred. No. 5,8e-194;
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTGACGAGGCGCCCAAGAGCAAGTGGCCCTGGCAGGTGAGCTGAGAG 76
 Db 66 GCATGCTGGGGGTGACGAGGCGCCCAAGAGCAAGTGGCCCTGGCAGGTGAGCTGAGAG 125
 QY 77 TCCAGGCGCCATTAATGAGTGAATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 136
 Db 126 TCCGGAACGATTAATGAGTGAATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGGTGC 185
 QY 137 TGACCCCGCGCGTGGTGGGAGACCGGACGTCGAAGATCTGGCCCGCCCTCAGGGTCAAC 196
 Db 186 TGAACCGGAGCGCACTGGTGGGAGACCGGACGTCGAAGATCTGGCCCGCCCTCAGGGTCAAC 245
 QY 197 TGGGGAGGACGACCTCTACTACAGGACCAAGCTGCTGGCGGTCAGAGGATCATGTGTGC 256
 Db 246 TCGGGGAGGACGACCTCTACTACAGGACCAAGCTGCTGGCGGTCAGAGGATCATGTGTGC 305
 QY 257 ACCCACAGTCTACACCCGCCAGATCGGAGCGGACATCGCCCTGCTGAGAGCTGAGAGG 316
 Db 306 ACCCACAGTCTCTACACCCGCCAGATCGGAGCGGACATCGCCCTGCTGAGAGCTGAGAGG 365
 QY 317 CGGTGAAGTCTCTCAAGCAGTCACACGCTGACCCCTGCCCTCCAGAGACCTTCC 376
 Db 366 CGGTGAAGTCTCTCAAGCAGTCACACGCTGACCCCTGCCCTCCAGAGACCTTCC 425
 QY 377 CCCCAGGAGATCGCTGCTGGGTCACTGGCTGGGGCGATGTGACATGATGAGCCCTCC 436
 Db 426 CCCCAGGAGATCGCTGCTGGGTCACTGGCTGGGGCGATGTGACATGATGAGCCCTCC 485
 QY 437 CACCGCATTTCTCTGAAACAGAGTGAAGGTCCCAATGAAACCAACATTTGTGACG 496
 Db 486 CACCGCATTTCTCTGAAACAGAGTGAAGGTCCCAATGAAACCAACATTTGTGACG 545
 QY 497 CAAAAATACACTTTGGCGCTTACACGGAGACGACGTCGCGATGTCCTGTGACGACATGC 556
 Db 546 CAAAAATACACTTTGGCGCTTACACGGAGACGACGTCGCGATGTCCTGTGACGACATGC 605
 QY 557 TGTGTGCGGGGAAACACCCCGGAGGAGCTCATGCGAGGGCGAATCCCGAAGGGCCCTGTGT 616
 Db 606 TGTGTGCGGGGAAACACCCCGGAGGAGCTCATGCGAGGGCGAATCCCGAAGGGCCCTGTGT 665
 QY 617 GCAAGGTGAATGAGCAGCTGCGTCAAGCGGGCGTGGTCACTGAGGGCGAGAGGCTGTGCC 676
 Db 666 GCAAGGTGAATGAGCAGCTGCGTCAAGCGGGCGTGGTCACTGAGGGCGAGAGGCTGTGCC 725
 QY 677 AGCCCAACCGGCGCTGGGATTAACACCGCGTGTACTTACTTGAATGAGCTGACCAACT 736
 Db 726 AGCCCAACCGGCGCTGGGATTAACACCGCGTGTACTTACTTGAATGAGCTGACCAACT 785
 QY 737 ATGTCCCAAAAAAGCCGTGAAGCGGCC 763
 Db 786 ATGTCCCAAAAAAGCCGTGAAGCGGCC 812

RESULT 15
 US-10-275-505-27

```

; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICATION: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAPALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Marian R.
; APPLICANT: AZIMZAI, Yalda; BULLIOTT, Vicki S.
; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Alna M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 27
; LENGTH: 2662
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7472460CB1
; US-10-275-505-27

```

```

Query Match          93.2%; Score 718.2; DB 18; Length 2662;
Beet Local Similarity 98.2%; Pred. No. 1.1e-191;
Matches 726; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 17 GAATGTCGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGGAGGTGAGCTGAGAG 76
DB 699 GATCGTTGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGGAGGTGAGCTGAGAG 758
QY 77 TCCAGGGCCCAATCTGATGCACTTCTGGGGGGCTCCCTCATCCACCCCAAGTGGTGC 136
DB 759 TCCGGGACCAGATACGATGCACTTCTGGGGGGCTCCCTCATCCACCCCAAGTGGTGC 818
QY 137 TGACCCGGGGGGGTGAGAGGCCCCCAGAGGCAAGTGGCCCTGGGAGGTGAGCTGAGAG 196
DB 819 TGACCCGAGCCAGCTGCGGGAGCCGAGGATGAGGATGGCCCGCTCAAGGGTGCAGAC 878
QY 197 TCGGGAGAGGAGCACTCTACTACTCAAGAGCAAGCTGCGGGTCAAGAGATCATCTGTC 256
DB 879 TCGGGAGAGGAGCACTCTACTACTCAAGAGCAAGCTGCGGGTCAAGAGATCATCTGTC 938
QY 257 ACCCAAGTTTCAACCCGCGGAGTGGAGCGGCAATGCGCTGCTGAGGAGAGAGC 316
DB 939 ACCCAAGTTTCAACCCGCGGAGTGGAGCGGCAATGCGCTGCTGAGGAGAGAGC 998
QY 317 CGATGAAGGTCTCAAGCAAGTCAAGCAAGGATCAAGCTGCGCCCTGAGGAGAGCCTTCC 376
DB 999 CGATGAAGGTCTCTCAAGCAAGTCAAGCAAGGATCAAGCTGCGCCCTGAGGAGAGCCTTCC 1058
QY 377 CCCCCGGGATGCGCGTGGTCACTGCGCTGGGGCGAGTGGACAATGATGAGCGCCTCC 436
DB 1059 CCCCCGGGATGCGCGTGGTCACTGCGCTGGGGCGAGTGGACAATGATGAGCGCCTCC 1118

```

```

QY 437 CACCGCATTTCTCTGAAGGAGTGAAGGTCCTCATTAATGAAAAACCAATTTGTGACG 496
DB 1119 CACCGCATTTCTCTGAAGGATGTGAAGTCCCAATTAATGAAAAACCAATTTGTGACG 1178
QY 497 CAAAATACCACTTTGGGCGCTTACACAGGAGAGCAAGTCCGATGTCCTGAGCAATGC 556
DB 1179 CAAAATACCACTTTGGGCGCTTACACAGGAGAGCAAGTCCGATGTCCTGAGCAATGC 1238
QY 557 TGTGTGCGGGGAAACACCGGAGGAGCTCATGSCAGGGGACTCCGAGAGGCGCTGGTGT 616
DB 1239 TGTGTGCGGGGAAACACCGGAGGAGCTCATGSCAGGGGACTCCGAGAGGCGCTGGTGT 1298
QY 617 GCAAGTGAATGGAGCACTGGCTGAGGGCGGGCGTGCAGGCTGGGGGAGGGGCTGTGCC 676
DB 1299 GCAAGTGAATGGAGCACTGGCTGAGGGCGGGCGTGCAGGCTGGGGGAGGGGCTGTGCC 1358
QY 677 AGCCCAACCGGCTGGGATCTACACCCCGTGTCACTTGAAGTGAATGCAAGCACT 736
DB 1359 AGCCCAACCGGCTGGGATCTACACCCCGTGTCACTTGAAGTGAATGCAAGCACT 1418
QY 737 ATGTCCCAAAAAGCCGTG 755
DB 1419 ATGTCCCAAAAAGCCGTG 1437

```

```

Search completed: August 27, 2005, 21:22:53
Job time : 630 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 18:58:19 ; Search time 188 Seconds
(without alignments)
6710.475 Million cell updates/sec

Title: US-09-598-982C-20

Perfect score: 1 999ccccctcgagaagaat.....cftcgaagcggcgcctcgt 771

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCRUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	764.6	99.2	771	3	US-09-079-970A-4
2	731	94.8	1128	2	US-09-016-366A-20
3	731	94.8	1128	2	US-08-978-404B-15
4	731	94.8	1137	2	US-09-016-366A-18
5	731	94.8	1137	2	US-08-978-404B-13
6	728.6	94.5	735	3	US-09-079-970A-1
7	726.2	94.2	1081	2	US-09-016-366A-22
8	726.2	94.2	1081	2	US-08-978-404B-17
9	726.2	94.2	1081	4	US-09-917-254-50
10	686.2	89.0	1154	2	US-09-016-366A-16
11	686.2	89.0	1154	2	US-08-978-404B-11
12	477.2	61.9	1219	2	US-08-978-404B-7
13	469.6	60.9	1108	2	US-09-016-366A-14
14	469.6	60.9	1108	2	US-08-978-404B-20
15	463	60.1	1031	2	US-08-978-404B-1
16	445.6	57.8	1103	2	US-09-016-366A-24
17	444	57.6	1097	2	US-08-978-404B-4
18	371	48.1	2259	3	US-08-845-998-3
19	371	48.1	2259	3	US-09-206-537-3
20	369.4	47.9	371	4	US-09-430-854-3
21	369.4	47.9	2218	2	US-08-845-998-5
22	369.4	47.9	2218	3	US-09-206-537-5
23	369.4	47.9	2218	3	US-09-430-854-5
24	271.8	25.3	1095	2	US-08-978-404B-9
25	201.4	26.1	980	4	US-09-023-942A-30
26	201.4	26.1	1110	3	US-09-386-653A-1
27	199.8	25.9	1212	4	US-09-620-312D-431

ALIGNMENTS

Result No.	Score	Query	Match Length	ID	Description
28	193.8	25.1	1130	3	US-09-386-653A-8
29	187.8	24.4	1613	4	US-09-387-375-1
30	187.8	24.4	1613	4	US-10-041-400A-1
31	187.8	24.4	1613	4	US-10-042-091A-1
32	181.2	23.5	1130	4	US-09-387-375-8
33	181.2	23.5	1130	4	US-10-041-400A-8
34	181.2	23.5	1130	4	US-10-042-091A-8
35	153	19.8	3757	2	US-09-016-366A-13
36	153	19.8	3757	2	US-08-978-404B-19
37	152	19.7	933	4	US-09-023-942A-29
38	143.8	18.7	2397	2	US-08-978-404B-2
39	140.6	18.2	1142	3	US-09-386-642-7
40	140.6	18.2	1142	3	US-09-386-642-8
41	132.2	17.1	1154	4	US-09-636-382A-1
42	131.2	17.0	1165	4	US-09-023-942A-28
43	131.2	17.0	1378	4	US-09-907-194A-262
44	131.2	17.0	1378	4	US-09-905-125A-262
45	131.2	17.0	1378	4	US-09-902-775A-262

RESULT 1
US-09-079-970A-4
Sequence 4, Application US/09079970A
Patent No. 6274366
GENERAL INFORMATION:
APPLICANT: Maffitt, Mark A.
APPLICANT: Niles, Andrew L.
APPLICANT: Haak-Frendscho, Mary
TITLE OF INVENTION: Enzymatically-Active Recombinant Human
TITLE OF INVENTION: Beta-Tryptase and Method of Making Same
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Intellectual Property Department
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WISCONSIN
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,970A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leone, Joseph T.
REGISTRATION NUMBER: 37,170
REFERENCE/DOCKET NUMBER: 34506.073
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 7..753
FEATURE:
NAME/KEY: misc_signal

LOCATION: 7.118
US-09-079-970A-4

Query Match 99.2%; Score 764.6; DB 3; Length 771;
Best Local Similarity 99.5%; Pred. No. 1.4e-169;
Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GGGCCCTCGAGAAAAGAAATCGTGGGGGGTCAAGAGAGCCCGCCGAGAGAAATGGCCCTGG 60
1 GGGCCCTCGAGAAAAGAAATCGTGGGGGGTCAAGAGAGCCCGCCGAGAGAAATGGCCCTGG 60
61 CAGGTAGCCCTGAGAGTCCAGCCGCGCCATCTGGATCACTTCTGCGGGGGCTCCCTCATC 120
61 CAGGTAGCCCTGAGAGTCCAGCCGCGCCATCTGGATCACTTCTGCGGGGGCTCCCTCATC 120
121 CACCCCAAGTGGTGTGACCCGAGCCGCACTGGTGGACCCGGACCTCAAGATCTGGCC 180
121 CACCCCAAGTGGTGTGACCCGAGCCGCACTGGTGGACCCGGACCTCAAGATCTGGCC 180
181 GCCCTCAGGGTGAACCTGGGGAGCACTCTACTAACAAGACCAAGCTGCTGCCGCTC 240
181 GCCCTCAGGGTGAACCTGGGGAGCACTCTACTAACAAGACCAAGCTGCTGCCGCTC 240
241 AGCAGGATCATCTGACCCAGCCAGCTTCTACACCGCCAGATCGGAGCAATCGCCCTG 300
241 AGCAGGATCATCTGACCCAGCCAGCTTCTACACCGCCAGATCGGAGCAATCGCCCTG 300
301 CTTGAGCTGAGAGAGCCGGTGAAGGTTCTCAGCCACTCAACCGCTCAACCTGCCCT 360
301 CTTGAGCTGAGAGAGCCGGTGAAGGTTCTCAGCCACTCAACCGCTCAACCTGCCCT 360
361 GCCTCAGAGACTTCCCGCCCGGGAGTCCGCTGGGTCACTGGTGGGGGAGATGGAC 420
361 GCCTCAGAGACTTCCCGCCCGGGAGTCCGCTGGGTCACTGGTGGGGGAGATGGAC 420
421 AATGATGAGCGCTCCCAACCGCATTTCTCTGAGACAGAGTGAAGTCCCATATGAA 480
421 AATGATGAGCGCTCCCAACCGCATTTCTCTGAGACAGAGTGAAGTCCCATATGAA 480
481 AACCAATTTTNGACGGAAATTCACACTTTGGGCTTACAACGGAGAAACAAGTCCGCAATC 540
481 AACCAATTTTNGACGGAAATTCACACTTTGGGCTTACAACGGAGAAACAAGTCCGCAATC 540
541 GTCCTGAGCAGACTGTGTGTGCGGGGAAACACCCGAGGAACTCATGCCAGGCGACTCC 600
541 GTCCTGAGCAGACTGTGTGTGCGGGGAAACACCCGAGGAACTCATGCCAGGCGACTCC 600
601 GGAAGGCCCCCTGTGTGCAAGGTGAATGGCACTGGCTGCAAGCGGGCTGTGAGCTGG 660
601 GGAAGGCCCCCTGTGTGCAAGGTGAATGGCACTGGCTGCAAGCGGGCTGTGAGCTGG 660
661 GGGAGAGGCTGTGCGCAGCCCAACCGGCTGGGATCTAACAACCGTCACTACTACTTG 720
661 GGGAGAGGCTGTGCGCAGCCCAACCGGCTGGGATCTAACAACCGTCACTACTACTTG 720
721 GACTGATCCACCACTATGTCCCAAAAAGCCGTGAAGCGCCGCTGT 771
721 GACTGATCCACCACTATGTCCCAAAAAGCCGTGTGAAGCGCGCTGTGT 771

RESULT 2
US-09-016-366A-20
; Sequence 20; Application us/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
; STREET: 600 Atlantic Avenue

CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1128 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-016-366A-20

Query Match 94.8%; Score 731; DB 2; Length 1128;
Best Local Similarity 98.7%; Pred. No. 1.1e-161;
Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

17 GAATCGTGGGGGTGAGAGAGCCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 76
17 GAATCGTGGGGGTGAGAGAGCCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 76
87 GATCGTGGGGGTGAGAGAGCCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 146
87 GATCGTGGGGGTGAGAGAGCCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 146
147 TCACAGCCCACTAATGATGACCTTCTGCGGGGCTCCCTGATCAACCCCAAGTGGTGC 206
147 TCACAGCCCACTAATGATGACCTTCTGCGGGGCTCCCTGATCAACCCCAAGTGGTGC 206
137 TGAACCGCCGCGCTGCGTGGAGCCGAGCAAGTCAAGATCTGGCCCTCAAGGGTGAAC 196
137 TGAACCGCCGCGCTGCGTGGAGCCGAGCAAGTCAAGATCTGGCCCTCAAGGGTGAAC 196
207 TGACCGCAGCGACTGCGTGGAGCCGAGCAAGTCAAGATCTGGCCCTCAAGGGTGAAC 266
207 TGACCGCAGCGACTGCGTGGAGCCGAGCAAGTCAAGATCTGGCCCTCAAGGGTGAAC 266
197 TGGGGAGCAGCACTCTACTAACAAGTCAAGCAAGTCTGCGGGTCAAGCAAGTCTGTC 256
197 TGGGGAGCAGCACTCTACTAACAAGTCAAGCAAGTCTGCGGGTCAAGCAAGTCTGTC 256
267 TGGGGAGCAGCACTCTACTAACAAGTCAAGCAAGTCTGCGGGTCAAGCAAGTCTGTC 326
267 TGGGGAGCAGCACTCTACTAACAAGTCAAGCAAGTCTGCGGGTCAAGCAAGTCTGTC 326
257 ACCCAAGTCTAACAACCCCGCAGATCGGAGCGGACATCGCCCTGAGGCTGAGGAGC 386
257 ACCCAAGTCTAACAACCCCGCAGATCGGAGCGGACATCGCCCTGAGGCTGAGGAGC 386
327 ACCCAAGTCTAACAACCCCGCAGATCGGAGCGGACATCGCCCTGAGGCTGAGGAGC 386
327 ACCCAAGTCTAACAACCCCGCAGATCGGAGCGGACATCGCCCTGAGGCTGAGGAGC 386
317 CGGTGAAGTCTCAAGCAAGTCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCC 376
317 CGGTGAAGTCTCAAGCAAGTCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCC 376
387 CGGTGAAGTCTCAAGCAAGTCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCC 446
387 CGGTGAAGTCTCAAGCAAGTCAACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCC 446
377 CCCCCGAGTGGCTGTGGTCACTGGCTGGGCGGATGGATGGAGCAATGATGAGCCCTCC 436
377 CCCCCGAGTGGCTGTGGTCACTGGCTGGGCGGATGGATGGAGCAATGATGAGCCCTCC 436
447 CCCCCGAGTGGCTGTGGTCACTGGCTGGGCGGATGGATGGAGCAATGATGAGCCCTCC 506
447 CCCCCGAGTGGCTGTGGTCACTGGCTGGGCGGATGGATGGAGCAATGATGAGCCCTCC 506
437 CACGGCCATTTCTGTGAAGCAAGTGAAGTCCCAATATGAAAAACAATTTGTGACG 496
437 CACGGCCATTTCTGTGAAGCAAGTGAAGTCCCAATATGAAAAACAATTTGTGACG 496
507 CACCCCAATTTCTGTGAAGCAAGTGAAGTCCCAATATGAAAAACAATTTGTGACG 566
507 CACCCCAATTTCTGTGAAGCAAGTGAAGTCCCAATATGAAAAACAATTTGTGACG 566
497 CAAAAATCAACCTTGGCCCTTACAGCGGAGCAAGTCTCGATCGTGAAGCAATGCG 556
497 CAAAAATCAACCTTGGCCCTTACAGCGGAGCAAGTCTCGATCGTGAAGCAATGCG 556
567 CAAAAATCAACCTTGGCCCTTACAGCGGAGCAAGTCTCGATCGTGAAGCAATGCG 626
567 CAAAAATCAACCTTGGCCCTTACAGCGGAGCAAGTCTCGATCGTGAAGCAATGCG 626

OY 557 TGTGTGCGGGAAACACCCGGAGGAGACTCATGCGAGGAGCCCTCGTGTGT 616
 DB 627 TGTGTGCGGGAAACACCCGGAGGAGACTCATGCGAGGAGCCCTCGTGTGT 686
 OY 617 GCAAGGTGAATGACACTGCTGTCAGGCGGCGGTGTGAGCTGGGGGAGGCTGTGCC 676
 DB 687 GCAAGGTGAATGACACTGCTGTCAGGCGGCGGTGTGAGCTGGGGGAGGCTGTGCC 746
 OY 677 AGCCCAACCGGCTGGCACTGACACCCGATGACCTACTGAGATGCAACCACT 736
 DB 747 AGCCCAACCGGCTGGCACTGACACCCGATGACCTACTGAGATGCAACCACT 806
 OY 737 ATGTCCCAAAAAGCCGTGAAGCGGCC 763
 DB 807 ATGTCCCAAAAAGCCGTGAAGCGGCC 833

RESULT 3

US-08-978-404B-15
 ; Sequence 15, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELETYPE:
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1128 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-978-404B-15
 Query Match 94.8%; Score 731; DB 2; Length 1128;
 Best Local Similarity 98.7%; Pred.No. 1,1e-161;
 Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 137 TGACCGCCGCGCGGTGCGGAGCCGGAAGCTCAAGATCTGGCCGCTCAGAGGTGAAC 196
 DB 207 TGACCGGAGGCACTGGGTGGAGCCGGAAGTCTGGCCGCTCAGAGGTGAAC 266
 OY 197 TGGGAGGACGACCTTACTACCAAGGACCAAGTGTGCGGTCAAGATCATGCTGC 256
 DB 267 TGGGAGGACGACCTTACTACCAAGGACCAAGTGTGCGGTCAAGATCATGCTGC 326
 OY 257 ACCGACGATTCATACAGCCGCAAGATGGAGCGGACATGCGCCGTGAGTGGAGG 316
 DB 327 ACCGACGATTCATACAGCCGCAAGATGGAGCGGACATGCGCCGTGAGTGGAGG 386
 OY 317 CCGTGAAGATCTCCAGGACATGCAAGGTCACCTGCCCCCTGCTCAGAGACTTCC 376
 DB 387 CCGTGAAGATCTCCAGGACATGCAAGGTCACCTGCCCCCTGCTCAGAGACTTCC 446
 OY 377 CCCCCGGGATGCGGTGCTGGTCACTGAGTGGGGGATGAGCAATGATGAGCGCTCC 436
 DB 447 CCCCCGGGATGCGGTGCTGGTCACTGAGTGGGGGATGAGCAATGATGAGCGCTCC 506
 OY 437 CACCGCATTTCTCTGAGACGAGTGAAGTCCCAATATGAAAAACAATTTGTGAGC 496
 DB 507 CACCGCATTTCTCTGAGACGAGTGAAGTCCCAATATGAAAAACAATTTGTGAGC 566
 OY 497 CAATAATCAACTTGGGCGCTACACGAGGAGACGCTCCGATGCTCCGTCAGACATGC 556
 DB 567 CAATAATCAACTTGGGCGCTACACGAGGAGACGCTCCGATGCTCCGTCAGACATGC 626
 OY 557 TGTGTGCGGGAAACACCCGGAGGAGACTCATGCGAGGCGACTCCGAGGCGCTGTGT 616
 DB 627 TGTGTGCGGGAAACACCCGGAGGAGACTCATGCGAGGCGACTCCGAGGCGCTGTGT 686
 OY 617 GCAAGGTGAATGACACTGCTGTCAGGCGGCGGTGTCACTGGGGGAGGCGTGTGCC 676
 DB 687 GCAAGGTGAATGACACTGCTGTCAGGCGGCGGTGTCACTGGGGGAGGCGTGTGCC 746
 OY 677 AGCCCAACCGGCTGGCACTGACACCCGATGACCTACTGAGATGCAACCACT 736
 DB 747 AGCCCAACCGGCTGGCACTGACACCCGATGACCTACTGAGATGCAACCACT 806
 OY 737 ATGTCCCAAAAAGCCGTGAAGCGGCC 763
 DB 807 ATGTCCCAAAAAGCCGTGAAGCGGCC 833

RESULT 4

US-09-016-366A-18
 ; Sequence 18, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,366A
 ; FILING DATE: January 30, 1998
 ; CLASSIFICATION: 530
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/037,090

FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-016-366A-18

Query Match 94.8%; Score 731; DB 2; Length 1137;
 Best Local Similarity 98.7%; Pred. No. 1,1e-161;
 Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

17 GAATGTCGGGGGCTCAGAGGGCCCGAGAGGCAAGTGGCCCTGGCAGGTCGAGAGG 76
 85 GCAATCGTGGGGGCTCAGAGGGCCCGAGAGGCAAGTGGCCCTGGCAGGTCGAGAGG 144
 77 TCCACGGCCCATACTGATGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGTGC 136
 145 TCCAGGGCCCATACTGATGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGTGC 204
 137 TGACCGCCGGGGCTGCGGTGGAGCCCGAAGTCAAGATCTGGCCGCTCGAGGGTGCAC 196
 205 TGACCGCCAGGCACTGCGGTGGAGCCCGAAGTCAAGATCTGGCCGCTCGAGGGTGCAC 264
 197 TGGGGAGGAGCACTCTACTACAGAGCAAGCTGCTGCGGGTCAAGAGATCATGTGC 256
 265 TGGGGAGGAGCACTCTACTACTACAGAGCAAGCTGCTGCGGGTCAAGAGATCATGTGC 324
 257 ACCCAGAGTTCTACACCGCCAGATCGAGCGGACATGCGCCCTGCTGAGCTGAGAGAC 316
 325 ACCCAGAGTTCTACACCGCCAGATCGAGCGGACATGCGCCCTGCTGAGCTGAGAGAC 384
 317 CGGTGAAGTCTTCAAGCAGTCAACGGTCAACCGTGCCTGCTGAGAGATCATGTGC 376
 385 CGGTGAAGTCTTCAAGCAGTCAACGGTCAACCGTGCCTGCTGAGAGATCATGTGC 444
 377 CCCGGGGAGTGGCGTCTGGGTCAGTGGCTGGGGGAGTGGAGCAATGATGAGCCGCTCC 436
 445 CCCGGGGAGTGGCGTCTGGGTCAGTGGCTGGGGGAGTGGAGCAATGATGAGCCGCTCC 504
 437 CACCGCATTTCTCTGAAAGCAGTGAAGTCCCTCCATAATGMAAACAATTTGTGACG 496
 505 CACCGCATTTCTCTGAAAGCAGTGAAGTCCCTCCATAATGMAAACAATTTGTGACG 564
 497 CAAAATACCACTTGGGGCTTACTACAGGAGAGAGAGTCCGGATGCTCCGTTGACAGATGC 556
 565 CAAAATACCACTTGGGGCTTACTACAGGAGAGAGAGTCCGGATGCTCCGTTGACAGATGC 624
 557 TGTGTCCCGGAAACACCCGAGAGGACTCATGACGAGGAGACTCCGAGAGGAGCCCTGTGT 616
 625 TGTGTCCCGGAAACACCCGAGAGGACTCATGACGAGGAGACTCCGAGAGGAGCCCTGTGT 684
 617 GCAAGGTGAATGGCACTTGGCTGCAAGCGGGGCTGTGTGACTGAGGGGCGAGGGCTGTGCC 676
 685 GCAAGGTGAATGGCACTTGGCTGCAAGCGGGGCTGTGTGACTGAGGGGCGAGGGCTGTGCC 744
 677 AGCCCAACCGGCTGGGATCTTACACCGGATGCACTTGGATGCACTGAGATCCAGACT 736
 745 AGCCCAACCGGCTGGGATCTTACACCGGATGCACTTGGATGCACTGAGATCCAGACT 804
 737 ATGTCCCAAAAAGCCGTGAAGGGCC 763
 805 ATGTCCCAAAAAGCCGTGAAGGGCC 831

RESULT 5
 US-08-978-404B-13
 ; Sequence 13, Application US/08978404B
 ; Patent No. 5668782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard J.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-978-404B-13

Query Match 94.8%; Score 731; DB 2; Length 1137;
 Best Local Similarity 98.7%; Pred. No. 1,1e-161;
 Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

17 GAATGTCGGGGGCTCAGAGGGCCCGAGAGGCAAGTGGCCCTGGCAGGTCGAGAGG 76
 85 GCAATCGTGGGGGCTCAGAGGGCCCGAGAGGCAAGTGGCCCTGGCAGGTCGAGAGG 144
 77 TCCACGGCCCATACTGATGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGTGC 136
 145 TCCAGGGCCCATACTGATGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCAAGTGGTGC 204
 137 TGACCGCCGGGGCTGCGGTGGAGCCCGAAGTCAAGATCTGGCCGCTCGAGGGTGCAC 196
 205 TGACCGCCAGGCACTGCGGTGGAGCCCGAAGTCAAGATCTGGCCGCTCGAGGGTGCAC 264
 197 TGGGGAGGAGCACTCTACTACAGAGCAAGCTGCTGCGGGTCAAGAGATCATGTGC 256
 265 TGGGGAGGAGCACTCTACTACTACAGAGCAAGCTGCTGCGGGTCAAGAGATCATGTGC 324
 257 ACCCAGAGTTCTACACCGCCAGATCGAGCGGACATGCGCCCTGCTGAGCTGAGAGAC 316
 325 ACCCAGAGTTCTACACCGCCAGATCGAGCGGACATGCGCCCTGCTGAGCTGAGAGAC 384
 317 CGGTGAAGTCTTCAAGCAGTCAACGGTCAACCGTGCCTGCTGAGAGATCATGTGC 376
 385 CGGTGAAGTCTTCAAGCAGTCAACGGTCAACCGTGCCTGCTGAGAGATCATGTGC 444

QY 377 CCCCCGGGGATGCGGCGTGGGCACTGGGCGGGGATGGGACAAATGATGAGGCGCCCTCC 436
 Db 445 CCCCCGGGGATGCGGCGTGGGCACTGGGCGGGGATGGGACAAATGATGAGGCGCCCTCC 504
 QY 437 CACCCTCCCTCTGAGAGCAGTGAAGTCCCTCAATATGAAAAACCACTTTGTGACG 496
 Db 505 CACCGCCATTTCTCTGAGAGCAGTGAAGTCCCTCAATATGAAAAACCACTTTGTGACG 564
 QY 497 CAAAATACCACTTTGGCGCTTACAGCGGAGACGATCCCGATCGTCCGTGACGACATGC 556
 Db 565 CAAAATACCACTTTGGCGCTTACAGCGGAGACGATCCCGATCGTCCGTGACGACATGC 624
 QY 557 TGTGTCGGGGAAACAACCGGAGAGGACTGATGCAAGGGGACCTCCGAGAGGGCCCTGGTGT 616
 Db 625 TGTGTCGGGGAAACAACCGGAGAGGACTGATGCAAGGGGACCTCCGAGAGGGCCCTGGTGT 684
 QY 617 GCAGAGTGAATGCACTGGCTGACAGCGCGGCGGTGTGATGAGTGGGCGAGGCTGTGCC 676
 Db 685 GCAGAGTGAATGCACTGGCTGACAGCGCGGCGGTGTGATGAGTGGGCGAGGCTGTGCC 744
 QY 677 AGCCCAACCGGCTGGCACTTACACCCGCTGCACTTACTTGGACTGGATCCACT 736
 Db 745 AGCCCAACCGGCTGGCACTTACACCCGCTGCACTTACTTGGACTGGATCCACT 804
 QY 737 ATGTCCCAAAAAAGCGGTGAAGCGGCGC_763
 Db 805 ATGTCCCAAAAAAGCGGTGAAGCGGCGC_831

RESULT 6

US-09-079-970A-1
 ; Sequence 1, Application US/09079970A
 ; Patent No. 6274366
 ; GENERAL INFORMATION:
 ; APPLICANT: Mafiltc, Mark A.
 ; APPLICANT: Niles, Andrew L.
 ; APPLICANT: Haak-Frendscho, Mary
 ; TITLE OF INVENTION: Enzymatically-Active Recombinant Human
 ; TITLE OF INVENTION: Beta-Trypsin and Method of Making Same
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Intellectual Property Department
 ; STREET: 8000 Excelsior Drive, Suite 401
 ; CITY: Madison
 ; STATE: WISCONSIN
 ; COUNTRY: U.S.A.
 ; ZIP: 53717-1914
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/079, 970A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Leone, Joseph T.
 ; REGISTRATION NUMBER: 37,170
 ; REFERENCE/DOCKET NUMBER: 34506, 073
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 831-2100
 ; TELEFAX: (608) 831-2106
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 735 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO

ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..735
 US-09-079-970A-1

Query Match 94.5%; Score 728, 6; DB 3; Length 735;
 Best Local Similarity 99.5%; Pred. No. 3,7e-161;
 Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 19 ATCTGCGGGGTACAGAGGCGCCCGGAGGAAAGTGGCCCTGGCAGAGTGAAGCTGAGAGTC 78
 Db 1 ATCTGCGGGGTACAGAGGCGCCCGGAGGAAAGTGGCCCTGGCAGAGTGAAGCTGAGAGTC 60
 QY 79 CAGGCGCCATTCAGAGTGCATTTGCGGGGCTCCCTCATTCACCCCGAGTGGTGTCTG 138
 Db 61 CAGGCGCCATTCAGAGTGCATTTGCGGGGCTCCCTCATTCACCCCGAGTGGTGTCTG 120
 QY 139 ACCGCGCGGGGTGCGGGGACCGGAGCTGCAAGGATCTGGCCCGCTCAGGGTGCACACTG 198
 Db 121 ACCGCGCGGGGTGCGGGGACCGGAGCTGCAAGGATCTGGCCCGCTCAGGGTGCACACTG 180
 QY 199 CGGAGGAGCACTCTTACTTACAGAGCAAGCTGCTGCGGCTCAGCAGATCATGTGCAC 258
 Db 181 CGGAGGAGCACTCTTACTTACAGAGCAAGCTGCTGCGGCTCAGCAGATCATGTGCAC 240
 QY 259 CCAAGATTTCTACACCGGCTCAGAGTGGAGCGGACATTCGCTGCTGGAGCTGGAGAGCG 318
 Db 241 CCAAGATTTCTACACCGGCTCAGAGTGGAGCGGACATTCGCTGCTGGAGCTGGAGAGCG 300
 QY 319 GTGAAGTCTCAGGCAAGTCCACAGGCTGACCCCTGCGCCCTGCTCAGAGACTTTCGCC 378
 Db 301 GTGAAGTCTCAGGCAAGTCCACAGGCTGACCCCTGCGCCCTGCTCAGAGACTTTCGCC 360
 QY 379 CCGGGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
 Db 361 CCGGGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 439 CCGCCATTTCTCTGAGAGCAGTGAAGTCCCAATATGAAAAACCACTTTGTGACGCA 498
 Db 421 CCGCCATTTCTCTGAGAGCAGTGAAGTCCCAATATGAAAAACCACTTTGTGACGCA 480
 QY 499 AAATACCACTTTGGCGCTTACAGCGGAGACAGGCTCCGCAATCGCTGGTGAACATGCTG 558
 Db 481 AAATACCACTTTGGCGCTTACAGCGGAGACAGGCTCCGCAATCGCTGGTGAACATGCTG 540
 QY 559 TGTCCCGGAAACAACCGGAGGACTCATGCGAGGGGACTCCGAGGGGCCCTGTGTGTC 618
 Db 541 TGTCCCGGAAACAACCGGAGGACTCATGCGAGGGGACTCCGAGGGGCCCTGTGTGTC 600
 QY 619 AAGGTGAATGCACTTGGCTGACAGCGCGGGGTGTGACTGCTGGGGGAGGGCTGTGCCAG 678
 Db 601 AAGGTGAATGCACTTGGCTGACAGCGCGGGGTGTGACTGCTGGGGGAGGGCTGTGCCAG 660
 QY 679 CCGAACCAGCGGCTGCACTTACACCCGCTGCACTTACTTGGACTGGATCCACCATAT 738
 Db 661 CCGAACCAGCGGCTGCACTTACACCCGCTGCACTTACTTGGACTGGATCCACCATAT 720
 QY 739 GTCCCAAAAAAGCGG 753
 Db 721 GTCCCAAAAAAGCGG 735

RESULT 7

US-09-016-366A-22
 ; Sequence 22, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Huang, Chifu
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS

NUMBER OF SEQUENCES: 65
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,366A
 FILING DATE: January 30, 1998
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,090
 FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1081 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 US-09-016-366A-22

Query Match 94.2%; Score 726.2; DB 2; Length 1081;
 Best Local Similarity 98.3%; Pred. No. 1.5e-160;
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

17 GAATCGTGGGGGGTCCGAGAGGCCCCCAGAGGAAAGTGGCCCTGGCAGGTGAAGCTGAGAG 76
 66 GCATCGTGGGGGTCCAGAGGCCCCCAGAGGAAAGTGGCCCTGGCAGGTGAAGCTGAGAG 125
 77 TCACAGCGCCCATCTGGATGACTTCTGCGGGGGCTCCCTCATTCACCCCAAGTGGGTGC 136
 126 TCGGCGACGGATCTGGATGCACTTCTGGGGGGCTCCCTCATTCACCCCAAGTGGGTGC 185
 137 TGACCGCCGGGGGTGGTGGGAGCCGGAAGTCTCAAGATCTGGCCCTTCAGGGGTCAAAC 196
 186 TGACCGGAGGGCACTGGTGGGAGCCGGAAGTCTCAAGATCTGGCCCTTCAGGGGTCAAAC 245
 197 TGAGGGAGGAGCACTCTACTATCACAGGACCGAGTGGTGGCCCGGTCAGAGGATGATGATG 256
 246 TGGGGAGGAGCACTCTACTATCACAGGACCGAGTGGTGGCCCGGTCAGAGGATGATGATG 305
 257 ACCACAGTCTTCAACCGCCAGATGCGAGCGGACATCGCCCTCTGAGCTTGAGAGC 316
 306 ACCACAGTCTTCAACCGCCAGATGCGAGCGGACATCGCCCTCTGAGCTTGAGAGC 365
 317 CGGTGAAGGTCTTCAAGCCAGTCCACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 376
 366 CGGTGAAGGTCTTCAAGCCAGTCCACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 425
 377 CCGGGGGAGTCCGGTGGTGGTCACTGGTGGGGGAGTGGGACATGATGATGATGATGATGATG 436
 426 CCGGGGGAGTCCGGTGGTGGTCACTGGTGGGGGAGTGGGACATGATGATGATGATGATGATG 485
 437 CACCGCATTTCTTCTGAGAGCAAGGTCCCAATTAATGAAAAACAATTGAGAG 496
 486 CACCGCATTTCTTCTGAGAGCAAGGTCCCAATTAATGAAAAACAATTGAGAG 545

497 CAAAATACCACTTGGCGGCTTACAGCGGAGAGCAAGTCCGATCGTCCGTGACGACATGC 556
 546 CAAAATACCACTTGGCGGCTTACAGCGGAGAGCAAGTCCGATCGTCCGTGACGACATGC 605
 557 TGTGTGCCGGGAAACACCCGGAGGGACTCATGACAGGGGACTCCGGAGGGCCCTGGGTGT 616
 606 TGTGTGCCGGGAAACACCCGGAGGACTCATGACAGGGGACTCCGGAGGGCCCTGGGTGT 665
 617 GCAAGTGAATGGCACTGGCTGCAGCGGGGCTGTCAGCTGGGGCGAGGGCTGTGCC 676
 666 GCAAGTGAATGGCACTGGCTGCAGCGGGGCTGTCAGCTGGGGCGAGGGCTGTGCC 725
 677 AGCCCAACCGCGCTGGCATCTACACCCGTGTCACTACTTGGAGTGTCACTCACT 736
 726 ACCCAACCGCGCTGGCATCTACACCCGTGTCACTACTTGGAGTGTCACTCACT 785
 737 ATGTCCCAAAAAGCGGTGAAGCGGSC 763
 786 ATGTCCCAAAAAGCGGTGAAGTCAAGGC 812

RESULT 8
 US-08-978-404B-17
 Sequence 17, Application US/089978404B
 Patent No. 5968782
 GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1081 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-978-404B-17

Query Match 94.2%; Score 726.2; DB 2; Length 1081;
 Best Local Similarity 98.3%; Pred. No. 1.5e-160;
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

17 GAATCGTGGGGGGTCCGAGAGGCCCCCAGAGGAAAGTGGCCCTGGCAGGTGAAGCTGAGAG 76
 66 GCATCGTGGGGGTCCAGAGGCCCCCAGAGGAAAGTGGCCCTGGCAGGTGAAGCTGAGAG 125

QY 77 TCACAGCCGCACTGATGATGACCTTCTGCGGGGGCTCCCTCATCCACCCCGATGGGTGC 136
 DB 126 TCCGGAACCGATACGATGATGACCTTCTGCGGGGGCTCCCTCATCCACCCCGATGGGTGC 185
 QY 137 TGACCCGCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 196
 DB 186 TGACCCGCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 245
 QY 197 TGCGGAGACGACCTCTAATCAACGAGACCAAGCTGCTGCGGTCAGGACGATCTGCTGC 256
 DB 246 TGCGGAGACGACCTCTAATCAACGAGACCAAGCTGCTGCGGTCAGGACGATCTGCTGC 305
 QY 257 ACCCAGATTCTACACCGCCGATCGGAGCGGACATGAGCCCTGCTGAGCTGAGAGAC 316
 DB 306 ACCCAGATTCTACACCGCCGATCGGAGCGGACATGAGCCCTGCTGAGCTGAGAGAC 365
 QY 317 CGGTGAAGGTCTCAAGCCAGTCACACGATGACCCCTGCGCTCAAGACCTTCC 376
 DB 366 CGGTGAAGGTCTCAAGCCAGTCACACGATGACCCCTGCGCTCAAGACCTTCC 425
 QY 377 CCGCGGGGATGCGGTCGTCGATCTGCGTGGGCGGATGTGACATATGATGAGCGCTTCC 436
 DB 426 CCGCGGGGATGCGGTCGTCGATCTGCGTGGGCGGATGTGACATATGATGAGCGCTTCC 485
 QY 437 CACCGCATTTCTCTGAAAGAGGTGAAAGGTCCCATATGAGAAACCAATTTGTGACG 496
 DB 486 CACCGCATTTCTCTGAAAGAGGTGAAAGGTCCCATATGAGAAACCAATTTGTGACG 545
 QY 497 CAAAATACCACTTTGCGGCTTACACGCGGAGACGACGTCGCTGCTGCTGACGATGC 556
 DB 546 CAAAATACCACTTTGCGGCTTACACGCGGAGACGACGTCGCTGCTGCTGACGATGC 605
 QY 557 TGTGTGCGGGGAAACCCGAGAGACTATGTCGCAAGGGGACTCCGGAAGGGCCCTGTGTG 616
 DB 606 TGTGTGCGGGGAAACCCGAGAGACTATGTCGCAAGGGGACTCCGGAAGGGCCCTGTGTG 665
 QY 617 GCAAGGTGAATGGCACTGAGTGCAGGGGCGGCTGATGAGTCCGAGGGGCTGTGCC 676
 DB 666 GCAAGGTGAATGGCACTGAGTGCAGGGGCGGCTGATGAGTCCGAGGGGCTGTGCC 725
 QY 677 AGCCCAACCGGCTGCGATCTACACCCGCTGTCACCTACTACTTGGACTGATCCACACT 736
 DB 726 AGCCCAACCGGCTGCGATCTACACCCGCTGTCACCTACTACTTGGACTGATCCACACT 785
 QY 737 ATGTCCCAAAAAGCCGTGAAGCGGCGC 763
 DB 786 ATGTCCCAAAAAGCCGTGAAGCGGCGC 812

RESULT 9
 US-09-917-254-50
 ; Sequence 50, Application US/09917254
 ; Patent No. 6703204
 ; GENERAL INFORMATION:
 ; APPLICANT: Mutter, George
 ; APPLICANT: Baak, Jan
 ; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
 ; FILE REFERENCE: B0801/724 (JRV)
 ; CURRENT APPLICATION NUMBER: US/09/917,254
 ; PRIORITY FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: US 60/222,093
 ; PRIORITY FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 50
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-09-917-254-50

Query Match 94.2%; Score 726.2; DB 4; Length 1081;
 Best Local Similarity 98.3%; Pred. No. 1.5e-160;
 Matches 734; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 GAATCGTGGGGGTCAAGAGAGCCGCCAGAGCAAGTGGCCCTGGCAGTGTAGGCTGAGAG 76
 DB 66 GCATCGTGGGGGTCAAGAGAGCCGCCAGAGCAAGTGGCCCTGGCAGTGTAGGCTGAGAG 125
 QY 77 TCACAGCCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 136
 DB 126 TGACCCGCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 185
 QY 137 TGACCCGCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 196
 DB 186 TGACCCGCGCGCGCTGCGTGGGACCCGACGTCGAAAGATCTGAGCCGCTCAAGGTGCAAC 245
 QY 197 TGCGGAGACGACCTCTAATCAACGAGACCAAGCTGCTGCGGTCAGGACGATCTGCTGC 256
 DB 246 TGCGGAGACGACCTCTAATCAACGAGACCAAGCTGCTGCGGTCAGGACGATCTGCTGC 305
 QY 257 ACCCAGATTCTACACCGCCGATCGGAGCGGACATGAGCCCTGCTGAGCTGAGAGAC 316
 DB 306 ACCCAGATTCTACACCGCCGATCGGAGCGGACATGAGCCCTGCTGAGCTGAGAGAC 365
 QY 317 CGGTGAAGGTCTCAAGCCAGTCACACGATGACCCCTGCGCTCAAGACCTTCC 376
 DB 366 CGGTGAAGGTCTCAAGCCAGTCACACGATGACCCCTGCGCTCAAGACCTTCC 425
 QY 377 CCGCGGGGATGCGGTCGTCGATCTGCGTGGGCGGATGTGACATATGATGAGCGCTTCC 436
 DB 426 CCGCGGGGATGCGGTCGTCGATCTGCGTGGGCGGATGTGACATATGATGAGCGCTTCC 485
 QY 437 CACCGCATTTCTCTGAAAGAGGTGAAAGGTCCCATATGAGAAACCAATTTGTGACG 496
 DB 486 CACCGCATTTCTCTGAAAGAGGTGAAAGGTCCCATATGAGAAACCAATTTGTGACG 545
 QY 497 CAAAATACCACTTTGCGGCTTACACGCGGAGACGACGTCGCTGCTGCTGACGATGC 556
 DB 546 CAAAATACCACTTTGCGGCTTACACGCGGAGACGACGTCGCTGCTGCTGACGATGC 605
 QY 557 TGTGTGCGGGGAAACCCGAGAGACTATGTCGCAAGGGGACTCCGGAAGGGCCCTGTGTG 616
 DB 606 TGTGTGCGGGGAAACCCGAGAGACTATGTCGCAAGGGGACTCCGGAAGGGCCCTGTGTG 665
 QY 617 GCAAGGTGAATGGCACTGAGTGCAGGGGCGGCTGATGAGTCCGAGGGGCTGTGCC 676
 DB 666 GCAAGGTGAATGGCACTGAGTGCAGGGGCGGCTGATGAGTCCGAGGGGCTGTGCC 725
 QY 677 AGCCCAACCGGCTGCGATCTACACCCGCTGTCACCTACTACTTGGACTGATCCACACT 736
 DB 726 AGCCCAACCGGCTGCGATCTACACCCGCTGTCACCTACTACTTGGACTGATCCACACT 785
 QY 737 ATGTCCCAAAAAGCCGTGAAGCGGCGC 763
 DB 786 ATGTCCCAAAAAGCCGTGAAGCGGCGC 812

RESULT 10
 US-09-016-366A-16
 ; Sequence 16, Application US/09016366A
 ; Patent No. 595431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Huang, Chifu
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette

```

? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/016.366A
? FILING DATE: January 30, 1998
? CLASSIFICATION: 530
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 60/037,090
? FILING DATE: 05-FEB-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Plumer, Elizabeth R.
? REGISTRATION NUMBER: 36,637
? REFERENCE/DOCKET NUMBER: B0801/7093
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-720-3500
? TELEFAX: 617-720-2441
? TELEX:
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1154 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-09-016-366A-16

```

```

Query Match      89.0%; Score 686.2; DB 2; Length 1154;
Best Local Similarity 94.3%; Pred. No. 3.3e-151;
Matches 712; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

```

QY 17 GAATCGTGGGGGTGAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGGAGCTTAGAG 76
DB 106 GATTCGGGGGGGTGAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGGAGCTTAGAG 165
QY 77 TCACGGCCCAATCTGGATGCACTTCTGGGGGGGGCTCCCTCATCCACCCCAAGTGGAGTGC 136
DB 166 TCCGCAACCAATCTGGATGCACTTCTGGGGGGGGCTCCCTCATCCACCCCAAGTGGAGTGC 225
QY 137 TGAACCGCGCGCGCGTGGGGAGCCCGGACGTCGAAGATCTGCGCCGCTCAGGGGTCAAC 196
DB 226 TGAACCGCGCGCGCGTGGGGAGCCCGGACGTCGAAGATCTGCGCCGCTCAGGGGTCAAC 285
QY 197 TGGGGAGGACGACCTCTACTACAGAGCAAGTGGTGGCCGCTCAGCAGGATCATGTGC 256
DB 286 TGGGGAGGACGACCTCTACTACAGAGCAAGTGGTGGCCGCTCAGCAGGATCATGTGC 345
QY 257 ACCCAGAGTTCTACACCGCCAGATGGAGCGGAGCAATGCGCCCTGGAGCTGGAGAGG 316
DB 346 ACCCAGAGTTCTACACCGCCAGATGGAGCGGAGCAATGCGCCCTGGAGCTGGAGAGG 405
QY 317 CGGTGAAGGTCTCCAGCAGTCACACCGGTCAACCGTCCCGCCCTGCTCAGAGACTTCC 376
DB 406 CGGTGAAGGTCTCCAGCAGTCACACCGGTCAACCGTCCCGCCCTGCTCAGAGACTTCC 465
QY 377 CCCCCGGAGTGGCGTGGTGTGATGCTGGGGCGATGTGGAATGATGAGCGGCTTCC 436
DB 466 CCCCCGGAGTGGCGTGGTGTGATGCTGGGGCGATGTGGAATGATGAGCGGCTTCC 525
QY 437 CACCGCCATTTCTCTGGAAGCAGGTGAGGTCCCAATTAATGAAAACCAATTTTGAAG 496
DB 526 CACCGCCATTTCTCTGGAAGCAGGTGAGGTCCCAATTAATGAAAACCAATTTTGAAG 585
QY 497 CAAAATACCACTTTGGCGCCTTACACCGGAGACGACGTCGCGATGTCGATGACGATCC 556
DB 586 CAAAATACCACTTTGGCGCCTTACACCGGAGACGACGTCGCGATGTCGATGACGATCC 645
QY 557 TGTGTCCCGGAAACACCCGGAGGAGTCAATGCCAGGGGCACTCCGGAGGAGGCGCTGTGT 616
DB 646 TGTGTCCCGGAAACACCCGGAGGAGTCAATGCCAGGGGCACTCCGGAGGAGGCGCTGTGT 705
QY 617 GCAAGGTGAATGAGCAGCTGGCTGAGGCGGGGGTGGTCAAGTGGGGGGGAGGGGCTGTGCC 676

```

```

DB 706 GCAAGGTGAATGAGCAGCTGGCTACAGCGCGGGCGGTGATGCTGGAGCAGAGGCTGTGCC 765
QY 677 AGCCCAACCGCGCTGGATCTACACCGGATGCACTTGGATGATGATCCACT 736
DB 766 AGCCCAACCGCGCTGGATCTACACCGGATGCACTTGGATGATGATCCACT 825
QY 737 ATGTCCCAAAAAAGCGGTGAAAGCGGCGCGCTGT 771
DB 826 ATGTCCCAAAAAAGCGGTGAAAGCGGCGCGCTGT 860

```

```

RESULT 11
US-08-978-404B-11
? Sequence 11, Application US/08978404B
? Patent No. 5968782
? GENERAL INFORMATION:
? APPLICANT: Stevens, Richard I.
? TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
? NUMBER OF INVENTIONS: 74
? NUMBER OF SEQUENCES: 74
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
? STREET: 600 Atlantic Avenue
? CITY: Boston
? STATE: MA
? COUNTRY: U.S.A.
? ZIP: 02210-2211
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/978,404B
? FILING DATE: 25-NOV-97
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 60/032,354
? FILING DATE: 04-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Plumer, Elizabeth R.
? REGISTRATION NUMBER: 36,637
? REFERENCE/DOCKET NUMBER: B0801/7090
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-720-3500
? TELEFAX: 617-720-2441

```

```

? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1154 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-08-978-404B-11
Query Match      89.0%; Score 686.2; DB 2; Length 1154;
Best Local Similarity 94.3%; Pred. No. 3.3e-151;
Matches 712; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```


DB 286 TCGGGAGAGAGACCTTCACTACAGAGACAGACTGTGCTGACAGGATCATCTGCG 345
 QY 257 ACCCAAGTTTACACCCCGCAGATCGAGCGGACATGTCCTGCTGAGCTGGAGAGC 316
 DB 346 ACCCAAGTTTACACTACAGACTGAGCGGAGATATCCCTGCTGGAGCTGGAGAGC 405
 QY 317 CGGTAAAGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGC 376
 DB 406 CGGTAAAGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGC 465
 QY 377 CCCCCGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
 DB 466 CCCCCGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
 QY 437 CACCCGATTTCTTGAAGCAGGTTGAGGTTCCCAATTAAGAAACCAATTTGTAGC 496
 DB 526 CACCCGATTTCTTGAAGCAGGTTGAGGTTCCCAATTAAGAAACCAATTTGTAGC 585
 QY 497 CAAAATACACCTTGGCCGCTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
 DB 586 CAAAATACACCTTGGCCGCTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645
 QY 557 TGTGTCCGGGAAACACCCGAGAGAGACTCAATGCAAGGAGAGAGAGAGAGAGAG 616
 DB 646 TGTGTCCGGGAAACACCCGAGAGAGACTCAATGCAAGGAGAGAGAGAGAGAGAG 705
 QY 617 GCAAGGTGAAATGCACTGTGCTGCAAGGCGGCGGCTGCTGCTGCTGCTGCTG 676
 DB 706 GCAAGGTGAAATGCACTGTGCTGCAAGGCGGCGGCTGCTGCTGCTGCTGCTG 765
 QY 677 AGCCCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
 DB 766 AGCCCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
 QY 737 ATGTCCCAAAAAGCCGTGAAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
 DB 826 ATGTCCCAAAAAGCCGTGAAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860

RESULT 12
 ; Sequence 7, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MASP CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1219 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-978-404B-7

Query Match 61.9%; Score 477.2; DB 2; Length 1219;
 Best Local Similarity 77.1%; Pred. No. 2.5e-102;
 Matches 581; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 6 CTTGAGAAAAGATTCCTGGGGGCTCAGAGAGCCCGCAGAGAGAGAGAGAGAGAGAG 65
 DB 260 CCAAGAGTGGGGGCTTGTGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319
 QY 66 GAGCCTGAGAGTCCAGCGCCCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 125
 DB 320 GAGCCTGCTGCTCAATGAAACCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
 QY 126 CCAAGTGGGTGCTGACCGCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 185
 DB 380 ACAAGTGGGTGCTGACCGCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439
 QY 186 CAGAGTGGCACTGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
 DB 440 CAGAGTGGCACTGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
 QY 246 GATCATGTTGCAACCAAGTTCTACACCGCCAGATGGAGAGAGAGAGAGAGAGAGAG 305
 DB 500 GATCATGCAACCAAGTTCTACACCGCCAGATGGAGAGAGAGAGAGAGAGAGAGAGAG 559
 QY 306 GCTGAG 365
 DB 560 GCTGAG 619
 QY 366 AGAGACTTTCCCGGGAG 425
 DB 620 AGAGACTTTCCCGGGAG 679
 QY 426 TGAAGGCTCCCAACCGCATTTCTGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 485
 DB 680 TGAAGGCTCCCAACCGCATTTCTGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
 QY 486 CATTGTGACCAAAAATCACTTGGCGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 545
 DB 740 GCTTTGTGACCTGAAAGTATCAAAAAGAGTCTTACAGAGAGAGAGAGAGAGAGAG 799
 QY 546 TGAAGCAGATGCTGTGTGCGGGAAACACCCGAGAGAGAGAGAGAGAGAGAGAGAG 605
 DB 800 AGAGCAGATGCTGTGTGCGGGAAACACCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
 QY 606 GCCCCTGCTGAG 665
 DB 860 ACCCTGCTGCTGCAAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919
 QY 666 GGGCTGTGCCAG 725
 DB 920 GGGCTGTGCTGTGCCAG 979
 QY 726 GATCCAGCAGATGATGCTCCCAAAAAGCGGTGAAGC 759
 DB 980 GATCCAGCAGATGATGCTCCCAAAAAGCGGTGAATC 1013

RESULT 13
 ; Sequence 14, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.

APPLICANT: Huang, Chifu
 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 TITLE OF INVENTION: INHIBITORS
 NUMBER OF SEQUENCES: 65
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,366A
 FILING DATE: January 30, 1998
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,090
 FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1108 base pairs
 TYPE: nucleic acid
 STANDARDS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-016-366A-14

Query Match 60.9%; Score 469.6; DB 2; Length 1108;
 Best Local Similarity 77.2%; Pred. No. 1.5e-100;
 Matches 571; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

17 GAATGTCGGGGGTCAAGAGGCGCCCGGCGAGGAGTGGCGGCGGCGGCTGAGAGG 76
 125 GCATCGGGGAGGAGCATGAGGCTTCTGAGAGTAAGGGCCCTGGCAGGTGAGCTGAGAGT 184
 77 TCACAGCGCCCATACTGATGACCTTTGGGGGCTTCCCTCATCCACCCCGAGTGGGTG 136
 185 TTAAATTTAACTACTGATGACATTTCTGGCGAGGCTCTCATCCACCCGAGTGGGTG 244
 137 TGACCGCGGGGGGTGGGGAGCCGAGCGTCAAGGATCGGGCGGCGGCGGCTGAGAGG 196
 245 TCACCTGGGAGACACTGTTGGGACCGGACATCAAAAAGCCGACAGCTTTCCGGGGTCA 304
 197 TGGGGAGACGACTTACTTACCAAGACCGAGTGGTCCGGGTCAAGAGGATCATGTCG 256
 305 TTCGTGAGGAGTACTATATATATGAGGAGCAGCTCTTTTGAACCGGAGTGTGTCG 364
 257 ACCGACAGTTTACACCGCCGAGATGGAGCGGACATCGCCCTGCTGAGAGTGGAGAGG 316
 365 ACCCGACATTTACACCGCGGAGGAGTGGGAGCAAGTGGCCCTGGAGGCTTGAAGTCC 424
 317 CGGTGAAGTCTCTCAGGCACGTCACACCGGTCAACCCCTGCGCCCTGAGAGACCTTCC 376
 425 CTGTGATGTCCTCCACCATTTCACCCCATATCTCCCTGGCCCTGAGAGACCTTCC 484
 377 CCGCGGAGTCCGCTGGTGGGTCACTGGCTGGGGCGATGTGACAATGATGAGAGCGCTCC 436
 485 CCGCGGAGATCATGCTGGGTGAGCAGGCTGGGGCGCATTTGAAVAVVAGACAGACCTTCC 544
 437 CACCGCATTTCTCTGAGAGCAGGTGAGGTCCCCATATGAGAAAACACATTTGTGACG 496

DB 545 CACCTCCCTTATCTCTGGAAGCAAGTGAAGTTCGCCATTGTGAAAAACAGCCGTGTGACC 604
 QY 497 CAAAATACCACTTGGCGGCTCACAGGGAGAGAGATGTCGTCGCTGACGACAGC 556
 DB 605 GGAAGTACCACTGCGCTTACACGGAGATGATTTTCCATTTGTCCATGTGGCAATGC 664
 QY 557 TGTGTCGGGGAACAACCCGAGAGGACTCATGCAAGGCGACTCCGAGAGGCGCCCTGAGT 616
 DB 665 TGTGTCTGGAATATACAGAGAGAGACTCTCTGCAAGGGGACTTCAAGGGGCGCCACTG 724
 QY 617 GCAAGTGAATGACACTTGGCTGCAAGGGCGGCTGACTGGGGCGAGGGCTTGGCC 676
 DB 725 GCAAAGTGAAGGCTACTGCGCTGCAAGGAGAGTGTGACTGGGGGTGAGGCGTCCGCA 784
 QY 677 AGCCCAACCGGCTGACATCTACCCCGTGTGCACTACTTGGACTGATGACCAACT 736
 DB 785 AGCCCAACAGCCTGGCATCTACACCAGGATGACATCTTACTTGAATCTGATCCACCGCT 844
 QY 737 ATGTCCCAAAAAAGCCGTGA 756
 DB 845 ATGTCCCTGAGCATTTCCCTGA 864

RESULT 14
 US-08-978-404B-20
 ; Sequence 20, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 TITLE OF INVENTION: FIBRINOGEN
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1108 base pairs
 TYPE: nucleic acid
 STANDARDS: single
 TOPOLOGY: linear
 US-08-978-404B-20

Query Match 60.9%; Score 469.6; DB 2; Length 1108;
 Best Local Similarity 77.2%; Pred. No. 1.5e-100;
 Matches 571; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

17 GAATGTCGGGGGTCAAGAGGCGCCCGGCGAGGAGTGGCGGCGGCGGCTGAGAGG 76

```

125 GCATGATGAGGAGACATGAGGGCTTCTGAGAGATAGGGCCCTGGGACGATGAGAT 184
QY 77 TCACGCGCCCAATCTGAGATGACATTTTGGGGGGCTCCCTCATCCACCCCGATGGGTGC 136
Db 185 TTAATAATTAACACTGATGATATCTTCTGGGGGGCTCTCTCATCCACCCGATGGGTGC 244
QY 137 TGACCGCGGGGGGTGGGGGACCGGAGGATGATGGGGCCCGCCGATGGGTGC 196
Db 245 TCACTGCGGACATCTGTGGGGACCGCACATCAAAAAGCCCAAGCTTTTCCGGGTGCAGC 304
QY 197 TCGGGAGGAGGACCTCTACTACAGAGACAGCTGTGGCGGGTCAAGAGATCATGTGC 256
Db 305 TTTGAGAGAGATATCTATCTATGAGGAGACAGCTCTCTCTTTTGAACCGGATCTGTGC 364
QY 257 ACCCACAGTTCTACACCGCCAGATGAGCGGACATGCGCCCTGTGAGGTGAGAGAC 316
Db 365 ACCCCCACTATTACAGGGCCGAGGGGTGGGCAAGCGTTCCTGCTGGAGCTTGAAGTTC 424
QY 317 CGGTGAAGGTCTCCAGCCACGTCACAGCGTCAAGCTGCCCCCTGCTGAGAGACCTTGC 376
Db 425 CTGTGATGTCCTCAACCCATATCCACCCATATCCCTGCCCCCTGCTGGAGACCTTGC 484
QY 377 CCCCCGGGATGCGCGTGTCTGATCACTGCTGGGGCGATGTGACATGATGAGCCCTCC 436
Db 485 CCCCCTGGGACATGCTGTGGGTGACAGGCTGGGGCGACATGATATGACAGCCCTCTCC 544
QY 437 CACCCGCAATTTCTCTGAAAGCAGGTGAAAGGTCCCAATTAATGAAAAACAATTTGTGAC 496
Db 545 CACCTCTTATCTCTGAAAGCAAGTAAAGGTTCCTCCATTTGTGAAAAACAACCTGTGTGAC 604
QY 497 CAAAATACCACTTGGGGGCTTACAGGGGAGAGAGACGTCGCAATGCTGCTGAGAGACATGC 556
Db 605 GGAATGACCACTGCGCTCTTACAGGGAGATATTTCCCATTTGTCATGATGAGATGC 664
QY 557 TGTGTGCCCCGGAACAACCCGAGGAGCTCATGCCAGGGGACCTCCGAGGGCCCTGTGTGC 616
Db 665 TGTGTGCTGGAATAATCAAGAGAGACTCTCGCAGGGCGATTTCAAGGGGGCCACATGTCT 724
QY 617 GCAAGGTGAATGAGCACTGTGCTGACAGGGCGGGGTGTGACTGTTGGGGCGAGGGCTGTGC 676
Db 725 GCAAGGTGAATGAGGATCTGTGCTGACAGGGCGGGGTGTGACTGTTGGGGCGAGGGCTGTGC 784
QY 677 AGCCCAACCGGCTGATCTTACACCGGATGACCCGATCACTTGAATGATGATCCACAT 736
Db 785 AGCCCAACCGCTGATCTTACACCGGATGACCCGATGACCCGATGACCCGATGACCCGCT 844
QY 737 ATGTCCCAAAAAAGCCGTGA 756
Db 845 ATGTCCCTGATGACATTTCTGA 864

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978, 404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-978-404B-1

Query Match 60.1%; Score 463; DB 2; Length 1031;
Best Local Similarity 76.4%; Pred. No. 5,1e-99;
Matches 568; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

```

```

14 AAAGATGCTGGGGGTGACGAGAGGCCCCAGAGAGAGTGGCCCTGGCAGGTGAGCTGA 73
Db 106 AAGGATTTGGGGGGGACAGAGGACATGGGAAACAAGTGGCCCTGGCAGGTGAGCTGC 165
QY 74 GAGTCCAGGCGCCATCTGATGATGACATTTGAGGGGGCTCCCTGATCCACCCAGTGGG 133
Db 166 GTCCAAATGACACTCTGATGATGATTTCTGCGGTGGCTCCCTGATCCACCAAGTGGG 225
QY 134 TGTGACCCCGCGGCGTGGGACCCGAGCTCAAGATCTGGCCCTGAGGTGC 193
Db 226 TGTCTACTGGCGGACACTGTGTGGGACCGGATGTTGGTCAACCAAGGTCAGAGTAC 285
QY 194 AACTGCGGAGACACTCTTACTACAGAGACAGCTGTGCGCTGCGGATCAAGATCATGC 253
Db 286 AGCTCCGTAACAGTACCTCTATTACATGACCACTGATGATCTGTGAGCCAGATCATCA 345
QY 254 TGACCCCAAGTCTTACACCGCCCAAGATCCGAGAGGAGATGCGCTGCTGAGAGG 313
Db 346 CACACCCGACTTCTACATGCTTCCAGATGAGGGGAGATTTGCCCTGTAATCTACAA 405
QY 314 AGCCGATGAGGTCTCCAGCCAGCTTCCACAGCTCAACCCCTGCTCAAGAGACT 373
Db 406 ACCCTGTGAATTTCTGATCTATGCTTCCACCCCTGCTCCCTCAAGAGACT 465
QY 374 TCCCGCGGAGATGCGCTGTGGTCACTGCTGGGGGAGATGAGCAATGAGAGCCGC 433
Db 466 TCCCTCAGGAAAGTGTGGGTGACAGGCTGGGATGATGATGATGATGATGAAACC 525
QY 434 TCCCAACCGGCAATTTCTGATGATGAGAGTGAAGTCCCATATAGAAAAACAATTTGTG 493
Db 526 TCCGCGCACATTTCTTTTGAAGAGGTGCAAGTTCCCATATGAAAAACAACCTTTGTG 585
QY 494 AGCAAAATACCACTTTGGCCCTTACAGGAGAGAGAGCTGCGGATGCTCCCTGAGAGCA 553
Db 586 ACTTGAATATCACAAAGTGTCTCATCAAGGTGACAAATGTCACATTTGTCGGAATGACA 645
QY 554 TGTGTGTGCGGGGAAACCCCGGAGGACTGATGCAAGGGGAGACTCCGAGAGCCCTGG 613
Db 646 TGTGTGTGCTGGGAAATGAAAGACATGACTCTCCAGGGGAGACTCCGAGAGCCCTGG 705
QY 614 TGTGAAAGTGAATGGGACCTGCTGACAGGCGGGGCTGATGATGATGATGATGATGATG 673
Db 706 TGTGAAAGTGAATGGGACCTGCTGACAGGCGGGGCTGATGATGATGATGATGATGATG 765
QY 674 CCAAGCCCAACCGGCTGATCTTACACCCGATGACCTTACTACTGATGATGATGATGATG 733

```

```

RESULT 15
US-08-978-404B-1
; Sequence 1, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAAT CELL PROTEASE THAT CLEAVES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wolf, Greenfield & Sacke, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

```

Db 766 CACAGCCCAACAGGCTGCGATCTACACCCGGGTCACTATTTACTTGGACTGGATCCACC 825
Qy 734 ACTATGTCCCCCAAAAAGCCGTGA 756
Db 826 ACTATGTCCCCCAAGACTTCTGA 848

Search completed: August 27, 2005, 19:37:45
Job time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:19:35 ; Search time 56 Seconds
(without alignments)
472.993 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1 MSLILLALPVLASRAVAAP.....IYTRVTVYLDWIMHHVPRKP 275
Sequence: 1512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

- 1: DIRT:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1470.5	97.3	274	2 A45754	tryptase (EC 3.4.2
2	1430	94.6	275	2 C35863	tryptase (EC 3.4.2
3	1411	93.3	275	2 A35863	tryptase (EC 3.4.2
4	1405	92.9	275	2 B35863	tryptase (EC 3.4.2
5	1134	75.0	273	2 A47246	tryptase (EC 3.4.2
6	1133	74.9	276	2 A38654	mast cell proteina
7	1131	74.8	275	2 A32410	tryptase (EC 3.4.2
8	1109.5	73.4	270	2 S56160	mast cell tryptase
9	1105	73.1	274	2 JC4171	tryptase (EC 3.4.2
10	999	66.1	237	2 S68702	tryptase (EC 3.4.2
11	856	56.6	230	2 I48685	mast cell proteina
12	686.5	45.4	269	2 B32410	mastocytoma protei
13	516	34.1	343	1 A57014	proctactin (EC 3.4.
14	468	31.0	625	1 KFHU1	coagulation factor
15	465	30.8	638	1 KOHUP	plasma kallikrein
16	461	30.5	638	1 KOMSP1	plasma kallikrein
17	459.5	30.4	271	1 ELRRT2	pancreatic elastas
18	457.5	30.3	245	1 KYROB	chymotrypsin (EC 3
19	454	30.0	435	2 A61545	plasma kallikrein
20	454	30.0	638	1 KORPL	plasma kallikrein
21	451.5	29.9	812	1 PLBO	plasma (EC 3.4.21
22	448	29.6	460	2 B61545	plasma (EC 3.4.21
23	447.5	29.6	269	2 B26823	pancreatic elastas
24	440.5	29.1	263	2 A31299	chymotrypsin (EC 3
25	439	29.0	810	1 PLHU	plasma (EC 3.4.21
26	438.5	29.0	263	2 A21195	chymotrypsin (EC 3
27	438.5	29.0	271	2 A25528	pancreatic elastas
28	434	28.7	263	1 KYRTB	chymotrypsin (EC 3
29	434	28.7	810	2 B30848	plasma (EC 3.4.21

QY	DB	Score	Length	Match	Indels	Gaps
30	429.5	28.4	4548	1	S00657	
31	429	28.4	266	2	ELPG	apoptoteicn(a) (EC
32	429	28.4	1524	2	T30337	pancreatic elastas
33	428.5	28.3	367	2	JB0104	polyprotein - Afri
34	426.5	28.2	269	2	C26823	testicular serine
35	425.5	28.1	245	1	PLMS	pancreatic elastas
36	425	28.1	245	1	KYBOA	plasma (EC 3.4.21
37	425	28.1	245	1	PLPG	chymotrypsin (EC 3
38	424	28.0	246	2	B25528	plasma (EC 3.4.21
39	423.5	28.0	436	2	UX0172	actrosin (EC 3.4.21
40	421.5	27.9	810	2	I46260	plasma (EC 3.4.21
41	420.5	27.8	263	2	S47537	chymotrypsin (EC 3
42	417.5	27.6	270	2	A29934	pancreatic elastas
43	417	27.6	264	2	I38136	chymotrypsin-like
44	415	27.4	421	2	S29599	actrosin (EC 3.4.21
45	414	27.4	786	1	A47547	serine proteinase

ALIGNMENTS

RESULT 1
A45754
tryptase (EC 3.4.21.59) alpha precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 08-Sep-1997
C:Accession: A45754; B37193
R:Miller, J.S.; Westin, E.H.; Schwartz, L.B.
J. Clin. Invest. 84, 1188-1195, 1989
A:Title: Cloning and characterization of complementary DNA for human tryptase.
A:Reference number: A45754; MUID:90009311; PMID:2677049
A:Accession: A45754
A:Molecule type: mRNA
A:Residues: 1-274 <MID>
A:Cross-references: GB:M30038
R:Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A:Title: Cloning and characterization of a second complementary DNA for human tryptase.
A:Reference number: A37193; MUID:90369005; PMID:2203827
A:Molecule type: mRNA
A:Residues: 1-274 <MID>
A:Cross-references: GB:M30038
A>Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 205-Pro

C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase; zymogen
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-30/Domain: activation peptide #status predicted <ACT>
F:31-274/Product: tryptase I #status predicted <MAT>
F:31-266/Domain: trypsin homology <TRY>
F:74,120,223/Active site: His, Asp, Ser #status predicted

Query Match 97.3%; Score 1470.5; DB 2; Length 274;
Best Local Similarity 97.8%; Pred. No. 3.5e-123;
Matches 269; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 MSLILLALPVLASRAVAAPVQALQDAGIVGQEARSKMPQVSRVDRYMMHRCG 60
DB 1 MSLILLALPVLASRAVAAPVQALQDAGIVGQEARSKMPQVSRVDRYMMHRCG 60
QY 61 GSLLHPQWVLTFAACCLGPDVDLTLRVQLAEQHLVYQDQLLPVSRIIVHPQFIITGGA 120
DB 61 GSLLHPQWVLTFAACCLGPDVDLTLRVQLAEQHLVYQDQLLPVSRIIVHPQFIITGGA 119
QY 121 DIALLLEEPVNISSRVHTVWLPPASSTFPFGMPCWVTGWGDVNDDEPLPPFPKQYKV 180
DB 120 DIALLLEEPVNISSRVHTVWLPPASSTFPFGMPCWVTGWGDVNDDEPLPPFPKQYKV 179
QY 181 PIMENHICDACKYHIGAYTGDDVRIIRDDMLCAAGNSQSDSCKSGSGGRLVCKVNGTWLQAG 240
DB 180 PIMENHICDACKYHIGAYTGDDVRIIRDDMLCAAGNSQSDSCKSGSGGRLVCKVNGTWLQAG 239
QY 241 VVSMDEGCAQPNRPGIYTRVTVYLDWIMHHVPRKP 275

Db 240 VVSMDEGCAQPNRPFGITRVTYLLDMWIMHYYVPKKP 274

RESULT 2
tryptase (EC 3.4.21.59) III precursor - human
A:Residues: 1-275 <VAN>
A:Cross-references: UNIPROT:Q15661; GB:M33494; NID:G3927804; PIDN:AC83172.1; PID:G339977

Query Match 94.6%; Score 1430; DB 2; Length 275;
Best Local Similarity 93.5%; Pred. No. 1.4e-119;
Matches 257; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

1 MSLLILALPVLASRAVAAPVQALQQAGIVGGEGEABRSKMPQVSLRVDRRMYMHFCG 60
1 MNLLLLALPVLASRAVAAPVQALQGVIGGGGEABRSKMPQVSLRVDRRMYMHFCG 60
61 GSLHPOWVLTAAHCGSPRVDKDLAALRVQLRBQHLTYDDQLLPVSRITLVHGFYTAQIGA 120
61 GSLHPOWVLTAAHCGSPRVDKDLAALRVQLRBQHLTYDDQLLPVSRITLVHGFYTAQIGA 120
121 DIALLELEBPVNSSRVHTVMLPPASSETFPPGMPGVTGMDVNDERLPPFPPLKQYKV 180
121 DIALLELEBPVNSSRVHTVMLPPASSETFPPGMPGVTGMDVNDERLPPFPPLKQYKV 180
181 PIMENHICDAKXHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGGSGRPVCKVNGTWLQAG 240
181 PIMENHICDAKXHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGGSGRPVCKVNGTWLQAG 240
241 VVSMDEGCAQPNRPFGITRVTYLLDMWIMHYYVPKKP 275
241 VVSMDEGCAQPNRPFGITRVTYLLDMWIMHYYVPKKP 275

RESULT 3
tryptase (EC 3.4.21.59) I precursor - human
A:Residues: 1-275 <VAN>
A:Cross-references: UNIPROT:Q15661; GB:M33494; NID:G3927804; PIDN:AC83172.1; PID:G339977

A:Molecule type: DNA
A:Residues: 1-275 <VAN>
A:Cross-references: UNIPROT:Q15661; GB:M33494; NID:G3927804; PIDN:AC83172.1; PID:G339977

Query Match 93.3%; Score 1411; DB 2; Length 275;
Best Local Similarity 92.4%; Pred. No. 6.8e-118;
Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

1 MSLLILALPVLASRAVAAPVQALQQAGIVGGEGEABRSKMPQVSLRVDRRMYMHFCG 60
1 MNLLLLALPVLASRAVAAPVQALQGVIGGGGEABRSKMPQVSLRVDRRMYMHFCG 60
61 GSLHPOWVLTAAHCGSPRVDKDLAALRVQLRBQHLTYDDQLLPVSRITLVHGFYTAQIGA 120
61 GSLHPOWVLTAAHCGSPRVDKDLAALRVQLRBQHLTYDDQLLPVSRITLVHGFYTAQIGA 120
121 DIALLELEBPVNSSRVHTVMLPPASSETFPPGMPGVTGMDVNDERLPPFPPLKQYKV 180
121 DIALLELEBPVNSSRVHTVMLPPASSETFPPGMPGVTGMDVNDERLPPFPPLKQYKV 180
181 PIMENHICDAKXHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGGSGRPVCKVNGTWLQAG 240
181 PIMENHICDAKXHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGGSGRPVCKVNGTWLQAG 240
241 VVSMDEGCAQPNRPFGITRVTYLLDMWIMHYYVPKKP 275
241 VVSMDEGCAQPNRPFGITRVTYLLDMWIMHYYVPKKP 275

RESULT 4
tryptase (EC 3.4.21.59) II precursor - human
A:Residues: 1-275 <VAN>
A:Cross-references: UNIPROT:Q15661; GB:M33494; NID:G3927804; PIDN:AC83172.1; PID:G339977

A:Cross-references: UNIPROT:P20231; GB:M33492; NID:9339982; PID:AAA36779.1; PID:9339983
 A:Note: residues 2-275 are derived from mRNA; residue one was inferred from the genomic
 R:Miller, J.S.; Moxley, G.; Schwartz, L.B.
 J:Clin. Invest. 86, 864-870, 1990
 A:Title: Cloning and characterization of a second complementary DNA for human trypsinase.
 A:Reference number: A37193; MUID:90369005; PMID:2203827
 A:Accession: A37193
 A:Molecule type: mRNA
 A:Residues: 1-275 <MBL>
 A:Cross-references: GB:M37488; NID:9179583; PID:AAA51843.1; PID:9179584
 R:Blom, T.; Hellman, L.
 Scand. J. Immunol. 37, 203-208, 1993
 A:Title: Characterization of a trypsin mRNA expressed in the human basophil cell line K
 A:Reference number: 159473; MUID:93166209; PMID:8434231
 A:Accession: 159473
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-275 <RES>
 A:Cross-references: GB:S5551; NID:9265666; PID:AA013876.1; PID:94261576
 A:Experimental source: basophil cell line KU812
 C:Genetics:
 A:Gene: GDB:TPS1
 A:Cross-references: GDB:125890; OMTM:191080
 A:Map position: 16pter-16qter
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase; zymogen
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-30/Domain: activation peptide #status predicted <ACT>
 F:31-275/Product: trypsin I #status predicted <MAT>
 F:31-267/Domain: trypsin homology <TRY>
 F:74,121,122/Active site: His, Asp, Ser #status predicted

Query Match 92.9%; Score 1405; DB 2; Length 275;
 Best Local Similarity 92.0%; Pred. No. 2.3e-117;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAVAAPVQALQOAGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 60
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 1 MMLKLLPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58

DB 61 GSLIHPPQWVLTAAHCCLGPDVDLALTRVQLREOHLYYQDQLPVSRIIVHPQFYIIQTGA 120
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 59 GSLIHPPQWVLTAAHCCLGPDVDLALTRVQLREOHLYYQDQLPVSRIIVHPQFYIIQTGA 118

QY 121 DIALLEBEPVNISSRVHTTMLPPASSTFPFGMPCWVTGMDVNDDEPLPPPLKQYKV 180
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 119 DIALLEBEPVNISSRVHTTMLPPASSTFPFGMPCWVTGMDVNDDEPLPPPLKQYKV 178

DB 181 PIMENHICDQAKYHLAGYTGDDVRIIRDDMLCAGNSQRDSCGSGSGLVCKVNGTMIQAG 240
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 179 PIMENHICDQAKYHLAGYTGDDVRIIRDDMLCAGNSQRDSCGSGSGLVCKVNGTMIQAG 238

QY 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHYVPK 273
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 239 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHYVPK 271

C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:29-265/Domain: trypsin homology <TRY>

Query Match 75.0%; Score 1134; DB 2; Length 273;
 Best Local Similarity 74.4%; Pred. No. 2.8e-93;
 Matches 203; Conservative 23; Mismatches 45; Indels 2; Gaps 1;

QY 1 MSLILLALPVLASRAVAAPVQALQOAGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 60
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 1 MMLKLLPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58

DB 61 GSLIHPPQWVLTAAHCCLGPDVDLALTRVQLREOHLYYQDQLPVSRIIVHPQFYIIQTGA 120
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 59 GSLIHPPQWVLTAAHCCLGPDVDLALTRVQLREOHLYYQDQLPVSRIIVHPQFYIIQTGA 118

QY 121 DIALLEBEPVNISSRVHTTMLPPASSTFPFGMPCWVTGMDVNDDEPLPPPLKQYKV 180
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 119 DIALLEBEPVNISSRVHTTMLPPASSTFPFGMPCWVTGMDVNDDEPLPPPLKQYKV 178

DB 181 PIMENHICDQAKYHLAGYTGDDVRIIRDDMLCAGNSQRDSCGSGSGLVCKVNGTMIQAG 240
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 179 PIMENHICDQAKYHLAGYTGDDVRIIRDDMLCAGNSQRDSCGSGSGLVCKVNGTMIQAG 238

QY 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHYVPK 273
 ||:|||||LPLVLSLVAHAPEP-AMRREGIVGGQBARPRKMPQVSLRVRDRYMMHFCG 58
 239 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHYVPK 271

RESULT 6
 A38654
 mast cell proteinase 6 (BC 3.4.21.-) precursor - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 21-Feb-1992 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
 C:Accession: A38654; B38654; D35646; I59478
 R:Reynolds, D.S.; Gurtley, D.S.; Austen, K.F.; Serafin, W.E.
 J. Biol. Chem. 266, 3847-3853, 1991
 A:Title: Cloning of the cDNA and gene of mouse mast cell proteinase-6. Transcription by pK
 A:Reference number: A38654; MUID:91139682; PMID:1995638
 A:Accession: A38654
 A:Molecule type: DNA
 A:Residues: 1-276 <REY>
 A:Cross-references: UNIPROT:P21845; GB:M57625; NID:9200506; PID:AAA39987.1; PID:9200507
 A:Note: the authors translated the codon CGC for residue 24 as Ala, GAG for residue 37 as
 Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly
 A:Accession: B38654
 A:Molecule type: mRNA
 A:Residues: 1-276 <RE2>
 A:Cross-references: GB:M57626; NID:9200508; PID:AAA39988.1; PID:9200509
 R:Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990
 A:Title: Different mouse mast cell populations express various combinations of at least 6
 A:Reference number: A35646; MUID:90222202; PMID:2326280
 A:Accession: D35646
 A:Molecule type: protein
 A:Residues: 32-54 <RE3>
 R:Huang, K.; Adirink, M.; Gohl, A.E.; Nilsson, G.; Aveskog, M.; Larsson, L.G.; Nilsson, I
 Scand. J. Immunol. 38, 359-367, 1993
 A:Title: Expression of a mast cell trypsin in the human monocytic cell lines U-937 and n
 A:Reference number: I59478; MUID:94023807; PMID:8210998
 A:Accession: I59478
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-276 <RES>
 A:Cross-references: GB:L31853; NID:9473480; PID:AAA39725.1; PID:9473481
 C:Genetics:
 A:Gene: MMCP-6
 A:Introns: 24/1; 79/2; 168/1; 222/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase; zymogen
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-31/Domain: activation peptide #status predicted <ACT>
 F:32-276/Product: mast cell proteinase 6 #status experimental <MAT>

F:32-268/Domain: trypsin homology <TRY>
F:75,122,225/Active site: His, Asp, Ser #status predicted

Query Match 74.9%; Score 1131; DB 2; Length 276;
Best Local Similarity 73.3%; Pred. No. 3,5e-93;
Matches 198; Conservative 28; Mismatches 42; Indels 2; Gaps 1;

QY 4 LLLALLPVLASRAVAAPVQALQAGIVGGQAPRPSKMPQVSLRVRDRYMMHFCSGLI 63
DB 7 LLLMALSLIASLIVYASAPRP--ANQVRGIVGGEHASESKMPWQVSLRFLKNTYMHIFCGSGLI 64
QY 64 IHQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTSADTA 123
DB 65 IHQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTSADTA 124
QY 124 LLELEBPVNISSRVHTVMLPRPASETFPPGMPGCVWTGMDVNDDEPLPPFPPLKQVAVPIM 183
DB 125 LLELEBPVNVSTHINRISLPRASETFPPGMPGCVWTGMDVNDDEPLPPFPPLKQVAVPIM 184
QY 184 ENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCKDSGGPILVCKVNGTWLQAGVVS 243
DB 185 ENSLCDRKRKHTGTYTGDDPFIYHDGMLCAGNTRRDSQQDSSGGPILVCKVNGTWLQAGVVS 244
QY 244 WDEGCAQPNRPGIYTRVYTYLDMIHRYVPK 273
DB 245 WDEGCAQPNRPGIYTRVYTYLDMIHRYVPE 274

RESULT 7

A32410
trypsaee (EC 3.4.21.59) precursor - dog
C/Species: Canis lupus familiaris (dog)
C/Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jul-2004
C/Accession: A32410
R:Vanderstijle, P.; Craik, C.S.; Nadel, J.A.; Gaughey, G.H.
Biochemistry 28, 4148-4155, 1989
A>Title: Molecular cloning of dog mast cell tryptase and a related protease: structural
A/Reference number: A32410; MUID:89352460; PMID:2504277
A/Accession: A32410
A/Molecule type: mRNA
A:Residues: 1-275 <VAN>
A:Cross-references: UNIPROT:P15944; GB:M24664; NID:q163982; PIDN:AAA30854.1; PID:q163983
C/Superfamily: trypsin homology
C/Keywords: hydrolase; serine proteinase; zymogen
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-30/Domain: activation peptide #status predicted <ACT>
F:31-267/Domain: trypsin homology <TRY>
F:74,121,224/Active site: His, Asp, Ser #status predicted

Query Match 74.8%; Score 1131; DB 2; Length 275;
Best Local Similarity 72.7%; Pred. No. 5.3e-93;
Matches 200; Conservative 30; Mismatches 45; Indels 0; Gaps 0;

QY 1 MSLLLALPVLASRAVAAPVQALQAGIVGGQAPRPSKMPQVSLRVRDRYMMHFCSG 60
DB 1 MSPPLVLTALALGSLVSPAPQALQVGVIGREAPGSKMPWQVSLRFLKNGQYMRHICG 60
QY 61 GSTIHPQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTS 120
DB 61 GSTIHPQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTS 120
QY 121 DIALLLEBPVNISSRVHTVMLPRPASETFPPGMPGCVWTGMDVNDDEPLPPFPPLKQVAV 180
DB 121 DIALLLEBPVNVSAHVQVPTLRPALQTPPTGTRPCWVTGMDVHNSGTRPLRPPFPPLKQVAV 180
QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCKDSGGPILVCKVNGTWLQAG 240
DB 181 PIVENMCDVQYHGLSTGSDVIVREDMMLCAGNSKSDSCQDSSGGPILVCKVNGTWLQAG 240
QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHRYVPKP 275
DB 241 VVSWGEGCALPFRPGIYTRVYTYLDMIHRYVPE 275

RESULT 8

S56160
mast cell tryptase precursor - Mongolian jird
C/Species: Meriones unguiculatus (Mongolian jird)
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C/Accession: S56160
R:Murakumo, Y.; Ide, H.; Itoh, H.; Tomita, M.; Kobayashi, T.; Maruyama, H.; Horii, Y.; Ne
Biochem. J. 309, 921-926, 1995
A>Title: Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil, Meriones ur
A/Reference number: S56160; MUID:95366971; PMID:7639711
A/Accession: S56160
A/Status: preliminary
A/Molecule type: mRNA
A:Residues: 1-270 <MUR>
A:Cross-references: UNIPROT:P50342; EMBL:D31789; NID:g517122; PIDN:BAA06598.1; PID:g51712;
C/Superfamily: trypsin homology
F:26-262/Domain: trypsin homology <TRY>

Query Match 73.4%; Score 1109.5; DB 2; Length 270;
Best Local Similarity 73.3%; Pred. No. 4.2e-91;
Matches 200; Conservative 22; Mismatches 46; Indels 5; Gaps 2;

QY 1 MSLLLALPVLASRAVAAPVQALQAGIVGGQAPRPSKMPQVSLRVRDRYMMHFCSG 60
DB 1 MKLLLLALP--LFSIMHNSP----LCQEMGIVGGQAPRPSKMPQVSLRANETVYRHFCSG 55
QY 61 GSTIHPQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTS 120
DB 56 GSTIHPQWVLTAAHCLGPRVQKDLATLRVQLRQNHLYYQDQLLPVSRITVHPQFYITQTS 115
QY 121 DIALLLEBPVNISSRVHTVMLPRPASETFPPGMPGCVWTGMDVNDDEPLPPFPPLKQVAV 180
DB 116 DIALLLEBPVNISSRVHTVMLPRPASETFPPGMPGCVWTGMDVNDDEPLPPFPPLKQVAV 175
QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCKDSGGPILVCKVNGTWLQAG 240
DB 176 PIVENMCDVQYHGLSTGSDVIVREDMMLCAGNSKSDSCQDSSGGPILVCKVNGTWLQAG 235
QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHRYVPK 273
DB 236 VVSWGEGCALPFRPGIYTRVYTYLDMIHRYVPE 268

RESULT 9

JC4171
trypsaee (EC 3.4.21.59) precursor - rat
N:Alaternate names: mast cell tryptase
C/Species: Rattus norvegicus (Norway rat)
C/Date: 27-Aug-1995 #sequence_revision 27-Oct-1995 #text_change 09-Jul-2004
C/Accession: JC4171
R:Ride, H.; Itoh, H.; Tomita, M.; Murakumo, Y.; Kobayashi, T.; Maruyama, H.; Osada, Y.; Ne
J. Biochem. 119, 210-215, 1995
A>Title: cDNA sequencing and expression of rat mast cell tryptase.
A/Reference number: JC4171; MUID:96015171; PMID:8537314
A/Accession: JC4171
A/Molecule type: mRNA
A:Residues: 1-274 <IDE>
A:Cross-references: UNIPROT:P50343; DBJ:D38455; NID:g556555; PIDN:BAA07486.1; PID:g55655;
C/Comment: This enzyme is basically specific for a connective tissue mast cell, it is up-
enzyme inhibitors.
C/Superfamily: trypsin homology
C/Keywords: glycoprotein; hydrolase; mast cell; serine proteinase; zymogen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-29/Domain: activation peptide #status predicted <ACT>
F:30-274/Product: mast cell tryptase #status predicted <MAT>
F:30-266/Domain: trypsin homology <TRY>
F:73,120,223/Active site: His, Asp, Ser #status predicted
F:131/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 73.1%; Score 1105; DB 2; Length 274;
Best Local Similarity 72.0%; Pred. No. 1.1e-90;

Matches 195; Conservative 31; Mismatches 43; Indels 2; Gaps 1;

QY 4 LLLALLPYLASSRAAYAAPVQALQOAGIVGGQEARSRKMPQVSLRVRDRYMMHFCGSSL 63
Db 5 LLLALLSPLASLAVNAPCPVK--QGVGIVGGREASBSKMPQVSLRFRKSFMMHFCGSSL 62
QY 64 IHPQWVLTAAHCLGPDVQDLATLRVQLREOHLYYQDQLLPVSRIIVHPQFYIOTGADIA 123
Db 63 IHPQWVLTAAHCLGPDVQDLATLRVQLREOHLYYQDQLLPVSRIIVHPQFYIOTGADIA 122
QY 124 LLELEBPVNISRRVHTVMTPASSETFRPMPQVTKGWDVNDDEPLRPPRFLKQVKVPIV 183
Db 123 LLELEBPVNISRRVHTVMTPASSETFRPMPQVTKGWDVNDDEPLRPPRFLKQVKVPIV 182
QY 184 ENHICDAKYNHGLGAYTGDDVRIIRDDMLCAGNSQRDSCCKDSSGGLVCKVNGTWDAGVVS 243
Db 183 ENSLCDRKRKHTGLYGGDDPFIYQDGMTCAGNTRSDSCQDSSGGLVCKVNGTWDAGVVS 242
QY 244 WDEGAQPRRPGIYTRVYVYLLMHHYVYDK 274
Db 243 WDEGAQPRRPGIYTRVYVYLLMHHYVYDK 273

RESULT 10

688702
tryptase (EC 3.4.21.59) - bovine (fragment)
C/Species: Bos primigenius taurus (catle)
C/Date: 25-Feb-1998 #sequence_revison 13-Mar-1998 #text_change 09-Jul-2004
C/Accession: S68702; S68676
R.Pallaoro, M.; Gambacorti, A.; Fiorucci, L.; Mignogna, G.; Barra, D.; Accolli, F.
Eur. J. Biochem. 237, 100-105, 1996
A>Title: cDNA cloning and primary structure of tryptase from bovine mast cells, and evid
A/Reference number: S68702; PMID:96203914; PMID:8620861
A/Molecule type: mRNA
A/Residues: 1-237 <PAL>
A/Cross-references: UNIPROT:Q29464; EMBL:X94982; NID:g1332446; PIDN:CAA64438.1; PID:g133
A/Accession: S68676
A/Molecule type: protein
A/Residues: TVGGQDAP,1-8;61-74;90-97;126-148;162-190;208-222;227-235 <PAM>
C/Superfamily: trypsin; trypsin homology
C/Keywords: glycoprotein; hydrolase; serine proteinase; zymogen
F/1-229/Domain: trypsin homology (fragment) <TRY>
F/36,83,177/Active site: His, Asp, Ser #status predicted
F/94,195/Binding site: carbohydrate (Aem) (covalent) #status predicted

Query Match 66.1%; Score 999; DB 2; Length 237;
Best Local Similarity 74.6%; Pred. No. 2,4e-81;
Matches 176; Conservative 22; Mismatches 38; Indels 0; Gaps 0;
QY 40 SKMPQVSLRVRDRYMMHFCGSSLIHPQWVLTAAHCLGPDVQDLATLRVQLREOHLYYQD 99
Db 2 SQMPQVSLRVRDRYMMHFCGSSLIHPQWVLTAAHCLGPDVQDPSYFRVQLREOHLYYQD 61
QY 100 QLLPVSRIIVHPQFYIOTGADIALELEBPVNISRRVHTVMTPASSETFRPMPQVTKG 159
Db 62 QLLPVSRIIVHPQFYIOTGADIALELEBPVNISRRVHTVMTPASSETFRPMPQVTKG 121
QY 160 WQDVNDDEPLRPPRFLKQVKVPIVMEHNIICDAKYNHGLGAYTGDDVRIIRDDMLCAGNSQRD 219
Db 122 WQDVNDDEPLRPPRFLKQVKVPIVMEHNIICDAKYNHGLGAYTGDDVRIIRDDMLCAGNSQRD 181
QY 220 CKDSSGGLVCKVNGTWDAGVVS WDEGAQPRRPGIYTRVYVYLLMHHYVYDK 275
Db 182 CKDSSGGLVCKVNGTWDAGVVS WDEGAQPRRPGIYTRVYVYLLMHHYVYDK 237
RESULT 11
148685
mast cell proteainase 6 (EC 3.4.21.-) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I48685; S43172

R:Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A>Title: Genes for mast-cell serine protease and their molecular evolution.
A/Reference number: I48684; MUID:95048582; PMID:7959952
A/Accession: I48685
A>Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-230 <RES>

A/Cross-references: UNIPROT:P21845; EMBL:X78542; NID:g468809; PIDN:CAA55288.1; PID:g46881
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase
F/32-230/Domain: trypsin homology #status atypical <TRY>
Query Match 56.6%; Score 856; DB 2; Length 230;
Best Local Similarity 69.3%; Pred. No. 1.2e-66;
Matches 151; Conservative 26; Mismatches 39; Indels 2; Gaps 1;

RESULT 12

832410
mastocytoma proteainase (EC 3.4.21.-) precursor - dog
C/Species: Canis lupus familiaris (dog)
C/Date: 12-Oct-1989 #sequence_revison 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: B32410
R.Vanderlidge, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A>Title: Molecular cloning of dog mast cell tryptase and a related protease: structural e
A/Reference number: B32410; MUID:89352460; PMID:2504277
A/Accession: B32410
A/Molecule type: mRNA
A/Residues: 1-269 <VAN>
A/Cross-references: UNIPROT:P19236; GB:M24655; NID:g163984; PID:g163985; GB:J02862
A/Note: the authors translated the codon ACG for residue 114 as Ser
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-269/Product: mastocytoma proteainase #status predicted <MAT>
F/20-259/Domain: trypsin homology <TRY>
F/66,116,217/Active site: His, Asp, Ser #status predicted

Query Match 45.4%; Score 686.5; DB 2; Length 269;
Best Local Similarity 52.0%; Pred. No. 1.7e-53;
Matches 132; Conservative 34; Mismatches 79; Indels 9; Gaps 4;
QY 28 OAGIVGGQEARSRKMPQVSLR---VDRYMMHFCGSSLIHPQWVLTAAHCLGPDVQDLA 84
Db 17 KVGIVGGCKVPAKRYRMPQVSLRFRKSFMMHFCGSSLIHPQWVLTAAHCLGPDVQDLA 76
QY 85 TLRVQLREOHLYYQDQLLPVSRIIVHPQFYIOTG--ADIALLELEBPVNISRRVHTVM 141
Db 77 TLRVQLREOHLYYQDQLLPVSRIIVHPQFYIOTG--ADIALLELEBPVNISRRVHTVM 136
QY 142 LPPASSETFRPMPQVTKGWDVNDDEPLRPPRFLKQVKVPIVMEHNIICDAKYNHGLGAYTGDD 201
Db 137 LPPASSETFRPMPQVTKGWDVNDDEPLRPPRFLKQVKVPIVMEHNIICDAKYNHGLGAYTGDD 194
QY 202 VRIIRDDMLCAGNSQRDSCCKDSSGGLVCKVNGTWDAGVVS WDEGAQPRRPGIYTRVT 261

Db 195 DEVIKQDMICAGSEGHDSQMDSGSGLPVCRWKMCTWIQVGVSWGXCQGY-NLPGVYARVT 253
 Qy 262 YYLDWTHHHVPPKP 275
 Db 254 SYVSWIHQHIPLSP 267

RESULT 13

prolactin (EC 3.4.21.-) precursor - human
 A:7014
 C:Species: Homo sapiens (man)
 C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
 C:Accession: A57014; A54866
 R:Yu, J.X.; Chao, L.; Chao, J.
 J. Biol. Chem. 270, 13483-13489, 1995
 A:Title: Molecular cloning, tissue-specific expression, and cellular localization of hum
 A:Reference number: A57014; MUID:95266644; PMID:7768952
 A:Accession: A57014
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-343 <RSS>
 A:Cross-references: UNIPROT:O16651; GB:I41351; NID:g862304; PIDN:AAQ41759.1; PID:g862305
 A:Experimental source: prostate were determined by protein sequencing
 A:Note: parts of this sequence were determined by protein sequencing
 R:Yu, J.X.; Chao, L.; Chao, J.
 J. Biol. Chem. 269, 18843-18848, 1994
 A:Title: Prolactin is a novel human serine proteinase from seminal fluid. Purification,
 A:Reference number: A54866; MUID:94308140; PMID:8034638
 A:Accession: A54866
 A:Molecule type: protein
 A:Residues: 45-64 <YUA>
 C:Genetics:
 A:Gene: GDB:PRSS8
 A:Cross-references: GDB:676446; OMIM:600823
 A:Map position: 16p11.2-16p11.2
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-44,45-343/Product: prolactin #status predicted <MAT>
 F:33-44/Domain: prolactin light chain #status predicted <CHL>
 F:45-343/Domain: prolactin heavy chain #status predicted <CHH>
 F:45-281/Domain: trypsin homology <TRY>
 F:323-341/Domain: transmembrane #status predicted <TMML>
 F:37-154,70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted
 F:85,134,238/Active site: His, Asp, Ser #status predicted
 F:159/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 34.1%; Score 516; DB 1; Length 343;
 Best Local Similarity 40.0%; Pred. No. 3,3e-38;
 Matches 110; Conservative 42; Mismatches 109; Indels 14; Gaps 8;

4 LLLALPVLASRAUAAVAPVQLOAGIYGGDEARSKMPWVSLKRVDRYMMHFGGSL 63
 18 LVLVGLIRSGTQEGKALDFGVAPQARITGGSSAVKGMFVQVSTIYEG---VHVGGSGL 74
 Qy 64 IHWVWVLTAAACSLGRPVKDLATLRVQDREGNL--YYQD-QLLPVSRIIVHPQFYIIQTGA 120
 Db 75 VSWQWVLTAAACSF-PSEHNKEAVELVKAHQDLSYSEDAKVSFLKDIIPRPSV--IQBS 131

Qy 121 --DIALLEPERVNISSRVNTVMLPRASFPFGMFCVWGVDVNDVLPRLPFLKQV 178
 Db 132 QGGIALLQLSRPTTFRSYTRPRTIGLPAANMSPFVNGHICTVYWGCHVAPVSVLLTRKPLQDL 191
 Qy 179 KVDIMENHIGDAKYHIGAYTVGDVRIIRDDMLCAG--NSQRDSCKDGSGSGLVCKYNGTW 236
 Db 192 EVPLISRETCNCLYNDLAKREVRNHVQEDMVAAGIYVGGKADKCGSDGGLSCVPEGLM 250
 Qy 237 IQAGVVSMEDEGCAQPNRPGIYTRVTVYLDWVHHYV 271
 Db 251 YLTGIYSWGDAGARRRPGVYTLASSVASYIGSKV 285

RESULT 14

coagulation factor XIa (EC 3.4.21.27) precursor [validated] - human
 KFHU1
 N:Alternate names: antihemophilic factor C; plasma thromboplastin antecedent
 C:Species: Homo sapiens (man)
 C:Date: 13-Aug-1986 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
 C:Accession: A27431; A00920; A37940
 R:Asakai, R.; Davie, E.W.; Chung, D.W.
 Biochemistry 26, 7221-7228, 1987
 A:Title: Organization of the gene for human factor XI.
 A:Reference number: A27431; MUID:88107663; PMID:2827746
 A:Accession: A27431
 A:Molecule type: DNA
 A:Residues: 1-625 <ASA>
 A:Cross-references: UNIPROT:P03951; GB:M18295
 A:Note: the sequence shown follows the authors' translation
 R:Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
 Biochemistry 25, 2417-2424, 1986
 A:Title: Amino acid sequence of human factor XI, a blood coagulation factor with four tar
 A:Reference number: A00920; MUID:86243360; PMID:3636155
 A:Accession: A00920
 A:Molecule type: mRNA
 A:Residues: 1-625 <FNU>
 A:Cross-references: GB:M13142; NID:g182832; PIDN:AAA52487.1; PID:g182833
 R:McMillen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2056-2060, 1991
 A:Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
 A:Reference number: A37940; MUID:91152017; PMID:1998667
 A:Accession: A37940
 A:Molecule type: protein
 A:Residues: 28-33,35-49, 'X', 51-55, 'X', 57-63, 70-75, 'X', 77-79, 107-109, 'X', 111-112, 132-139, '
 ;280-282, 'X', 284, 285-297, 313-316, 'X', 318-319, 320-326; 'X', 328-330; 'X', 347-349; 373, 'X', 375;
 C:Comment: The proenzyme consists of two identical chains linked by one or more disulfide
 he active site, and a heavy chain, which associates with high molecular weight (HMW) kin
 C:Genetics:
 A:Gene: GDB:PI1
 A:Cross-references: GDB:119891; OMIM:264900
 A:Map position: 4q35-q435
 A:Introns: 19/1, 73/2, 109/1, 162/2, 199/1, 252/2, 289/1, 343/2, 379/1, 435/2, 494/1, 526
 C:Function:
 A:Description: catalyzes the proteolytic activation of coagulation factor IX
 A:Pathway: blood coagulation intrinsic pathway
 C:Superfamily: coagulation factor XI; trypsin homology
 C:Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-387/Product: coagulation factor XIa heavy chain #status experimental <HCH>
 F:19-108/Domain: apple repeat <AP1>
 F:109-198/Domain: apple repeat <AP2>
 F:199-288/Domain: apple repeat <AP3>
 F:290-379/Domain: apple repeat <AP4>
 F:388-625/Product: coagulation factor XIa light chain #status experimental <LCH>
 F:388-618/Domain: trypsin homology <TRY>
 F:20-103,514-581,571-599/Disulfide bonds: #status predicted
 F:29/Disulfide bonds: interchain #status experimental
 F:46-76,50-56,110-193,136-165,140-166,200-283,228-255,230-236,291-374,317-346,321-327,381
 F:90,126,353,450/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:339/Disulfide bonds: interchain #status predicted
 F:387-388/Cleavage site: Arg-11e (coagulation factor XIa) #status experimental
 F:331,480,575/Active site: His, Asp, Ser #status predicted
 F:491/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 31.0%; Score 468; DB 1; Length 625;
 Best Local Similarity 39.5%; Pred. No. 1.2e-33;
 Matches 98; Conservative 31; Mismatches 91; Indels 28; Gaps 6;

31 IYGGDEARSKMPWVSLKRVDRYMMHFGGSLIHPQVLTAAHCTGPDVVDLATTLYVQ- 89
 Db 388 IYGGTASVRSWMPVTVLHTTSPGRLHCGGSIINQWVILTAHCF-YGVEPRKILKRYVS 446
 Qy 90 --LRQHLYYQDPLPVRRIIVHPQFYIIQTGADTALLEPERVNISSRVNTVMLPRPASE 147
 Db 447 GILNQSSEIKEDTSFFGVQGEIITHDQYKMAESGVDIALLKLEFTVYVTSQSRICLPSRSGD 506

QY 148 TRPPGMPGCVTSMG-----DVNDNDEPLPPPPPLKQVKVPIVIMENHI CDAYKHLGAYTGD 201
 DB 507 RRVVITDTCVWTGMYGRKLRDKIQN-----TLQKAKIPLVYTNBECQKRYR----- 550
 QY 202 VRIIRDMLCAGNSQ--RDSCKGDGGPPLVCKVNGTWTLAGVVSMDGCAQPNRPGIYTR 259
 DB 551 GHKITHKMI CAGYREGKDACCKDGGGFLSCGKNEVWHLVGITSMGEGCAQSERPGVYTN 610
 QY 260 VTYTYLDWI 267
 DB 611 VVEYVDWI 618

RESULT 15

KOHUP
 Plasma kallikrein (EC 3.4.21.34) precursor - human
 N/Alternate names: kininogenin; plasma prekallikrein
 C/Species: Homo sapiens (man)
 C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 09-Jul-2004
 C/Accession: A00921; A37939
 R/Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
 Biochemistry 25, 2410-2417, 1986
 A/Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four
 A/Reference number: A00921; PMID:86243359; PMID:3521732
 A/Accession: A00921
 A/Molecule type: mRNA
 A/Residues: 1-638 <CHU>
 A/Cross-References: UNIPROT:P03952; GB:M13143; NID:G190262; PIDN:AAA60153.1; PID:G190263
 R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2050-2056, 1991
 A/Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of
 A/Reference number: A37939; PMID:91152016; PMID:19986866
 A/Accession: A37939
 A/Molecule type: protein
 A/Residues: 20-27;40-46; 'X',48;'H',50;'X',52-70;'H',75-76;'X',78-80;103-113;131-140;141-
 ;260-283;'X',285;287,291;'X',293-295;314-317;'X',319-320;321-324;'X',329-333;334-339;'X',
 525;528-551;562;'X',564-567;573;'X',575-576;578-583;'X',585;592-604 <MCM>
 C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
 C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 1
 are linked by one or more disulfide bonds.
 C/Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r
 inogen and may also play a role in the renin-angiotensin system by converting prorenin I
 C/Genetics:
 A/Gene: GDB:KLRK3
 A/Cross-References: GDB:127575; OMIM:229000
 A/Map position: 4q35-4q35
 C/Superfamily: coagulation factor XI; trypsin homology
 C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
 F.1-19/Domain: signal sequence #status predicted <SIG>
 F.20-638/Product: plasma kallikrein #status predicted <MAT>
 F.20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
 F.20-109/Domain: apple repeat <AP1>
 F.110-199/Domain: apple repeat <AP2>
 F.200-289/Domain: apple repeat <AP3>
 F.291-380/Domain: apple repeat <AP4>
 F.391-638/Domain: plasma kallikrein light chain #status predicted <LCH>
 F.391-621/Domain: trypsin homology <TRY>
 F.127-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383
 F.127,308,396,453,494/Binding site: carboxylate (Asn) (covalent) #status experimental
 F.318-347,340-345/Disulfide bonds: #status predicted
 F.390-391/Cleavage site: Arg-1le (coagulation factor XIIa) #status predicted
 F.434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 30.8%; Score 465; DB 1; Length 638;
 Best Local Similarity 39.0%; Pred. No. 2.3e-33;
 Matches 97; Conservative 41; Mismatches 81; Indels 30; Gaps 8;

QY 31 IYGGGQEARPSKWPQVSLAVRDRYWMHFGGSLIHPQWVLTAAHCL-GPDVNDLAIATRYQ 89
 DB 391 IYGGTNSWGEWPMQVSLQVKKLTAQRHLGSGSLIGHQVWLTAAHCFDGLPLQDV--WRIV 448
 QY 90 ---LREQHLYYQDQLPVSRIIVHPQFYITQCADIALLEPWNISSRVHTVMLPPAS 146

DB 449 SGILNLSIDITDTPFSQIKETIIHQNVKVSBNHNDIALIKLQADPLNLYTEPQKICLPLSPGK 508
 QY 147 ETRPPGMPGCVTSMG-----GDVNDNDEPLPPPPPLKQVKVPIVIMENHI CDAYKHLGAYTGD 200
 DB 509 DSTIITYTNCVWTGMYGRKLRDKIQN-----TLQKAKIPLVYTNBECQKRYR----- 553
 QY 201 VRIIRDMLCAGNSQ--RDSCKGDGGPPLVCKVNGTWTLAGVVSMDGCAQPNRPGIYTR 258
 DB 554 GHKITHKMI CAGYREGKDACCKDGGGFLSCGKNEVWHLVGITSMGEGCAQSERPGVYTN 610
 QY 260 VTYTYLDWI 267
 DB 611 VVEYVDWI 618

Search completed: August 27, 2005, 23:30:12
 Job time : 58 secs

This page is blank (photo)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:13:05 ; Search time 128 seconds
(without alignments)
1100.170 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1512
Sequence: 1 MLSLLLLALPVLASRAVAAP.....ITYRVTYVLDWIMHHVPRKP 275

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03: *
1: uniprot_sprot: *
2: uniprot_crembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	99.5	275	1 TRYA_HUMAN	P15157 homo sapien
2	1494	98.8	275	2 Q6CTM8	Q6ctm8 homo sapien
3	1494	98.8	276	2 Q6CUD8	Q6cud8 homo sapien
4	1414	93.5	275	2 Q6B051	Q6b051 homo sapien
5	1411	93.3	275	1 TRB1_HUMAN	Q15661 homo sapien
6	1405	92.9	275	1 TRB2_HUMAN	P20231 homo sapien
7	1405	92.9	282	2 Q6NZY1	Q6nzy1 homo sapien
8	1403	92.8	275	2 Q6FHB8	Q6fhb8 homo sapien
9	1401	92.7	275	2 Q6FRZ6	Q6frz6 homo sapien
10	1156	76.5	275	2 Q7YS62	Q7ys62 equus caball
11	1134	75.0	273	1 MCT7_MOUSE	Q02844 mus musculu
12	1133	74.9	276	1 MCT6_MOUSE	P21845 mus musculu
13	1131	74.8	275	1 TRYT_CANFA	P15944 canis famli
14	1123	74.3	273	2 Q9Z1N4	Q9z1n4 mus musculu
15	1120	74.1	233	2 Q96RZ7	Q96rz7 homo sapien
16	1118	73.9	242	1 TRYT_SHEEP	Q9xkm2 ovis aries
17	1110	73.4	273	2 Q6NMT8	Q6nbt8 homo sapien
18	1109.5	73.4	270	1 TRYT_MERUN	P50342 meriones un
19	1109	73.3	275	1 TRYT_PIG	Q9nzd1 sus scrofa
20	1105	73.1	273	1 MCT7_PAT	P27475 rattus norv
21	1105	73.1	273	2 Q6P6W8	Q6p6w8 rattus norv
22	1105	73.1	274	2 MCT6_PAT	P50343 rattus norv
23	1103	72.9	273	2 Q9XSM1	Q9xsm1 ovis aries
24	1100	72.8	235	1 TRYD_HUMAN	Q9bzj3 homo sapien
25	1098	72.6	271	2 Q84HY2	Q84hj2 bos taurus
26	999	66.1	237	2 Q29464	Q29464 bos taurus
27	717.5	47.5	280	2 Q66NX6	Q66nx6 canis famli
28	716.5	47.4	280	2 Q66NX5	Q66nx5 canis famli
29	697	46.1	318	2 Q80UR4	Q80ur4 mus musculu
30	692	45.8	314	2 Q6RUT2	Q6rut2 mus musculu
31	689	45.6	277	2 Q8SQ44	Q8sq44 sus scrofa

Result No.	Score	Query Match	Length	ID	Description
32	686.5	45.4	269	1 TRYM_CANFA	P19236 canis famli
33	649	42.9	279	2 Q99MS4	Q99ms4 mus musculu
34	632.5	41.8	321	1 TRYG_HUMAN	Q9nr2 homo sapien
35	628.5	41.6	321	2 Q96RZ6	Q96rz6 homo sapien
36	621	41.1	274	2 Q924N9	Q924n9 mus musculu
37	603.5	39.9	311	1 TRYG_MOUSE	Q9qu17 mus musculu
38	600.5	39.7	277	2 Q80WM7	Q80wm7 mus musculu
39	599.5	39.6	328	2 Q6BEA2	Q6bea2 rattus norv
40	597	39.5	321	2 Q61E60	Q61e60 rattus norv
41	592.5	39.2	328	2 Q8BUR6	Q8bur6 mus musculu
42	578	38.2	311	2 Q80XZ3	Q80xz3 rattus norv
43	576	38.1	330	2 Q6NVR7	Q6nvr7 xenopus tro
44	572.5	37.9	290	1 PR27_HUMAN	Q9bzq3 homo sapien
45	561.5	37.1	719	2 Q6DJ50	Q6dj50 xenopus tro

ALIGNMENTS

RESULT 1
 TRYA_HUMAN STANDARD; PRT; 275 AA.
 AC P15157; Q9HZY5; Q9UQI1;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Alpha-tryptase precursor (EC 3.4.21.59) (Trypsin 1).
 GN Name=TPSI;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lung;
 RA MEDLINE=90009311; PubMed=2677049;
 RA Miller J.S., Westin E.H., Schwartz L.B.;
 RT "Cloning and characterization of complementary DNA for human
 RT trypsinase.";
 RL J. Clin. Invest. 84:1188-1195(1989).
 RN [2]
 RP REVISIONS TO 89-93 AND 108.
 RA Schwartz L.B.;
 RN Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC MEDLINE=99121069; PubMed=9920877; DOI=10.1074/jbc.274.6.3355;
 RA Pallao M., Fejzo M.S., Shyesteh L., Blount J.L., Caughey G.H.;
 RT "Characterization of genes encoding known and novel human mast cell
 RT trypsinases on chromosome 16p13.3.";
 RL J. Biol. Chem. 274:3355-3362(1999).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Lung;
 RA Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
 RA Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
 RA Hunt J.E.;
 RT "Molecular cloning and characterization of novel human trypsinase cDNAs
 RT and splicing variants.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 31-50, AND PITUITARY, SEQUENCE OF 31-38.
 RC TISSUE=Lung;
 RA MEDLINE=87109258; PubMed=3543004;
 RA Cromlish J.A., Seidah N.G., Marcinkiewicz M., Hamelin J., Johnson D.A.,
 RA Chretien M.;
 RT "Human pituitary trypsinase: molecular forms, NH2-terminal sequence,
 RT immunocytochemical localization, and specificity with prothrombin and
 RT fluorogenic substrates.";
 RL J. Biol. Chem. 262:1363-1373(1987).
 CC -1- FUNCTION: Trypsinase is the major neutral protease present in mast
 CC cells and is secreted upon the coupled activation-degranulation
 CC response of this cell type.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
 CC with more restricted specificity than trypsin.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- SUBCELLULAR LOCATION: Released from the secretory granules upon
 CC mast cell activation.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P15157-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P15157-2; Sequence=VSP_005374;
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Trypsin
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M30038; AAA66934.1; -;
 DR EMBL; AF098328; AAD17846.1; -;
 DR EMBL; AF206665; AAG35695.1; -;
 DR EMBL; AF206666; AAG35696.1; -;
 DR PDB; 1LTD; X-ray; A/B/C/D=31-275.
 DR MEROPS; S01.143; -;
 DR Genew; HGNC:12018; TPST1.
 DR MIM; 191080; -;
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
 DR GO; GO:0006952; P:defense response; TAS.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00689; Trypsin_1.
 DR FRINGS; PRO0722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYPSPIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSPIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSPIN_SER; 1.
 DR 3d-structure; Alternative splicing; Direct protein sequencing;
 DR Glycoepitome; Hydrolyase; Polymorphism; Serine protease; Signal;
 KW Zymogen.
 KW Zymogen.
 FT SIGNAL 1 18 Potential.
 FT PROPEP 19 30 Activation peptide.
 FT CHAIN 31 275 Alpha-tryptase.
 FT ACT_SITE 74 74 Charge relay system.
 FT ACT_SITE 121 121 Charge relay system.
 FT ACT_SITE 224 224 Charge relay system.
 FT DISULFID 59 75
 FT DISULFID 155 230
 FT DISULFID 188 211
 FT DISULFID 220 248
 FT CARBOHYD 132 132 N-linked (GlcNAc...)(Potential).
 FT CARBOHYD 233 233 N-linked (GlcNAc...)(Potential).
 FT VARSPPLIC 79 87 Missing (in isoform 2).
 FT VARIANT 15 15 R->P (in alpha-II).
 FT VARIANT 221 221 R->Q (in alpha-II; dbSNP:1137382).
 FT CONFLICT 215 216 /Frid=VAR_012103.
 FT SEQUENCE 275 AA; 30772 MW; B9BAC4B8CB91CE75 CRC64;
 Query Match 99.5%; Score 1505; DB 1; Length 275;
 Best Local Similarity 99.3%; Pred.No.1.6e-122;
 Matches 273; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 61 GSLHPQWLTAAACGSPDYKDLATLTVQLREHLYYODDLIPVSRITIVHPQFYITQGA 120
 DB 61 GSLHPQWLTAAACGSPDYKDLATLTVQLREHLYYODDLIPVSRITIVHPQFYITQGA 120
 QY 121 DIALLELEPPVNISSRVHTVWLPSPSEFPMPMPVYVWGDVDDPPLPPPKQYKV 180
 DB 121 DIALLELEPPVNISSRVHTVWLPSPSEFPMPMPVYVWGDVDDPPLPPPKQYKV 180
 QY 181 PIMENHICDAKTHLGAYGGDVRITRDMLCAGNSORDSCGHSQSPVYCKVNGTWTQAG 240
 DB 181 PIMENHICDAKTHLGAYGGDVRITRDMLCAGNSORDSCGHSQSPVYCKVNGTWTQAG 240
 QY 241 VVSWDEGCAQENRPGIYRVTYVYLDVIMHHYVPKKP 275
 DB 241 VVSWDEGCAQENRPGIYRVTYVYLDVIMHHYVPKKP 275

RESULT 2
 Q86TW8 PRELIMINARY; PRT; 275 AA.
 AC Q86TW8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE TPST1 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Butler N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hallyk S.W.,
 RA Villalón D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC EMBL; BC028059; AAH28059.1; -;
 DR HSPF; P15157; ILTD.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00089; Trypsin_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYPSPIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSPIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 275 AA; 30713 MW; 189335BBCB473FA3 CRC64;

Query Match 98.8%; Score 1494; DB 2; Length 275;
Best Local Similarity 98.5%; Pred. No. 1.4e-121;
Matches 271; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLSLLLLALPYLASRAYAAPAPVQALQOAGIVGGQEA PRSKMPQVSLRVRDRYMMHFCG 60
DB 1 MLSLLLLALPYLASRAYAAPAPVQALQOAGIVGGQEA PRSKMPQVSLRVRDRYMMHFCG 60
OY 61 GSLIHPQWVLTAAHCLGPDVDLALTRVQLREQHLYYQDQLLPSRSRIIVHPQFYIIQTGA 120
DB 61 GSLIHPQWVLTAAHCLGPDVDLALTRVQLREQHLYYQDQLLPSRSRIIVHPQFYIIQTGA 120
OY 121 DIALLEEPPNISRVRHTVMWLPASSTFPFGMPCWVTGMDVNDDEPLPFPPLKQVKV 180
DB 121 DIALLEEPPNISRVRHTVMWLPASSTFPFGMPCWVTGMDVNDDEPLPFPPLKQVKV 180
OY 181 PIMENHICDAKYHGLGAYTGDVRIIRDDMLCAGNSORDSCKDSGGPLVCYKVNQTWLQAG 240
DB 181 PIMENHICDAKYHGLGAYTGDVRIIRDDMLCAGNTRRSDSCGDSGGPLVCYKVNQTWLQAG 240
OY 241 VVSWDEGCAQPNRPGIYTRVYVYLDWIIHHYVPKKP 275
DB 241 VVSWDEGCAQPNRPGIYTRVYVYLDWIIHHYVPKKP 275

RESULT 3

OY Q86UAS PRELIMINARY; PRT; 276 AA.

ID Q86UAS5

DT 01-JUN-2003 (TRMBLrel. 24, Created)

DT 01-JUN-2003 (TRMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TRMBLrel. 26, Last annotation update)

DE TP51 protein (Fragment).

GN Name=TP51;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLIN=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

Altechni S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci R.D., Mullaly S.J.,

Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalón D.K., Muzny D.M., Sodergren E., Madan A., Rodriguez S., Sanchez A.,

Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

DR HSSP; P15157; ILT0.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR003314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolyase; Protease; Serine protease.
KM Non_Ter
SQ SEQUENCE 276 AA; 30869 MW; 8CDC5EA037955F9C CRC64;
Query Match 98.8%; Score 1494; DB 2; Length 276;
Best Local Similarity 98.5%; Pred. No. 1.4e-121;
Matches 271; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLSLLLLALPYLASRAYAAPAPVQALQOAGIVGGQEA PRSKMPQVSLRVRDRYMMHFCG 60
DB 2 MLSLLLLALPYLASRAYAAPAPVQALQOAGIVGGQEA PRSKMPQVSLRVRDRYMMHFCG 61
OY 61 GSLIHPQWVLTAAHCLGPDVDLALTRVQLREQHLYYQDQLLPSRSRIIVHPQFYIIQTGA 120
DB 61 GSLIHPQWVLTAAHCLGPDVDLALTRVQLREQHLYYQDQLLPSRSRIIVHPQFYIIQTGA 121
OY 121 DIALLEEPPNISRVRHTVMWLPASSTFPFGMPCWVTGMDVNDDEPLPFPPLKQVKV 180
DB 121 DIALLEEPPNISRVRHTVMWLPASSTFPFGMPCWVTGMDVNDDEPLPFPPLKQVKV 181
OY 181 PIMENHICDAKYHGLGAYTGDVRIIRDDMLCAGNSORDSCKDSGGPLVCYKVNQTWLQAG 240
DB 181 PIMENHICDAKYHGLGAYTGDVRIIRDDMLCAGNTRRSDSCGDSGGPLVCYKVNQTWLQAG 241
OY 241 VVSWDEGCAQPNRPGIYTRVYVYLDWIIHHYVPKKP 275
DB 241 VVSWDEGCAQPNRPGIYTRVYVYLDWIIHHYVPKKP 276

RESULT 4

OY Q8B051 PRELIMINARY; PRT; 275 AA.

ID Q8B051

DT 25-OCT-2004 (TRMBLrel. 28, Created)

DT 25-OCT-2004 (TRMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TRMBLrel. 28, Last annotation update)

DE Beta tryptase 1.

GN Name=TPSBA1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pooled tissue;

RA MEDLIN=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

Altechni S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci R.D., Mullaly S.J.,

Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalón D.K., Muzny D.M., Sodergren E., Madan A., Rodriguez S., Sanchez A.,

Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Skimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smillius D.E., Schermerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Poiled tissue;
 RA Director MGC Project;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL: BC074975; AAH74975.1; -;
 DR GO: GO:0004263; F:chymotrypsin activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase; Protease; Serine protease.
 SQ SEQUENCE 275 AA; 30488 MW; A8F0B12C58F88C12 CRC64;

Query Match 93.5%; Score 1414; DB 2; Length 275;
 Best Local Similarity 92.7%; Pred. No. 1.3e-114;
 Matches 255; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MSLLILALPVLASRAVYAAPAVQALDQAGIVGGDEAPRSKRWQVSLRVRDRYMMHFCG 60
 Db 1 MSLLILALPVLASRAVYAAPAVQALDQAGIVGGDEAPRSKRWQVSLRVRDRYMMHFCG 60
 QY 61 GSLLHPQWVLTAAHCTGPRVVDLAVLREQHLVYDQDLIPVSKIIYHPQYIIQTGA 120
 Db 61 GSLLHPQWVLTAAHCTGPRVVDLAVLREQHLVYDQDLIPVSKIIYHPQYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASERTPRGMPGCVTGWGVNDEPRLPPRPKQKVY 180
 Db 121 DIALLELEBPVNISSRVHTVMLPPASERTPRGMPGCVTGWGVNDEPRLPPRPKQKVY 180
 QY 181 PIMENHIQAKYKHLGAYTGDVRIITRDMICAGNSQRDSCKSGDGGPLVCKVNGTWLQAG 240
 Db 181 PIMENHIQAKYKHLGAYTGDVRIITRDMICAGNSQRDSCKSGDGGPLVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPNRPFGIYTRVYTYLDMITHYVPRKP 275
 Db 241 VVSWDEGCAQPNRPFGIYTRVYTYLDMITHYVPRKP 275

RESULT 5
 TRBL_HUMAN STANDARD; PRT; 275 AA.
 ID TRBL_HUMAN Q15661; Q9H2Y4;
 AC Q15661; Q15663; Q9H2Y4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Trypsin beta-1 precursor (BC 3.4.21.59) (Trypsin 1) (Trypsin I).
 GN Name=TPSB1;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90251647; PubMed=2187193;
 RA Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
 RA Caughey G.H.;
 RT "Human mast cell trypase: multiple cDNAs and genes reveal a multigene

RT serine protease family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99121069; PubMed=9920877; DOI=10.1074/jbc.274.6.3355;
 RA Pallao M., Pejor M.S., Shyasteh L., Blount J.L., Caughey G.H.;
 RT "Characterization of genes encoding known and novel human mast cell
 RT trypases on chromosome 16p13.3.";
 RL J. Biol. Chem. 274:3355-3362(1999).
 RN [3]
 RP SEQUENCE OF 54-275 FROM N.A. (ISORF08 2).
 RA Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
 RA Heltzerich A., King G., Heywood G.O., Huang C., Stevens R.L.,
 RA Hunt J.E.;
 RT "Molecular cloning and characterization of novel human trypase cDNAs
 RT and splicing variants.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Trypsin is the major neutral protease present in mast
 CC cells and is secreted upon the coupled activation-degranulation
 CC response of this cell type.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|- , but
 CC with more restricted specificity than trypsin.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Released from the secretory granules upon
 CC mast cell activation.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q15661-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q15661-2; Sequence=VSP_005375;

CC -1- SIMILARITY: Belongs to the peptidase S1 family. Trypsase
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: M33494; AAC83172.1; -;
 DR EMBL: M33491; AA36778.1; -;
 DR EMBL: AF099144; AAD17860.1; -;
 DR EMBL: AF206667; AAC35697.1; -;
 DR PIR: A35863; A35863.
 DR HSSP: P20231; 1A0L.
 DR MEROPS: S01.015; -;
 DR MEROPS: S01.242; -;
 DR GeneW; HGNC:12019; TPSB1.
 DR MIM; 191081; -;
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Alternative splicing; Glycoprotein; HydroLase; Polymorphism;
 KW Serine protease; Signal; Zymogen.
 FT SIGNAL 1 18 Potential.
 FT PROPEP 19 30 Activation peptide (By similarity).
 FT CHAIN 31 275 Trypsin beta-1.
 FT ACT_SITE 74 74 Charge relay system.
 FT ACT_SITE 121 121 Charge relay system.
 FT ACT_SITE 224 224 Charge relay system.
 FT DISULFID 59 75 By similarity.
 FT DISULFID 155 230 By similarity.
 FT DISULFID 188 211 By similarity.
 FT DISULFID 220 248 By similarity.

-1- SIMILARITY: Belongs to the peptidase S1 family. Trypsase subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M37488; AAAS1843.1; -;
 DR EMBL: M34492; AAA36779.1; -;
 DR EMBL: M33493; AAA36780.1; -;
 DR EMBL: S55551; AAD13876.1; -;
 DR EMBL: AF099143; AAD17859.2; -;
 DR EMBL: AF099145; AAD17857.1; -;
 DR EMBL: AF099146; AAD17858.1; -;
 DR EMBL: BC029356; AAH29356.1; -;
 DR PIR: B35863; B35863.
 DR PDB: 1A0L; X-ray; A/B/C/D=31-274.
 DR PDB: 1AAO; Model; @=31-275.
 DR MEROPS: S01.242; -; -;
 DR GeneW; HGNC:14120; TP5B2.
 DR H-InvDB; HIX0012676; -;
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:004252; F:serine-type endopeptidase activity; NAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS50240; TRYPESIN_DOM; 1.
 DR PROSITE; PS00134; TRYPESIN_HIS; 1.
 DR PROSITE; PS00135; TRYPESIN_SER; 1.
 DR 3D-structure; Glycoprotein; Hydrolase; Polymorphism; Serine protease;
 KW Signal; Zymogen.
 FT FT SIGNAL 1 18 Potential.
 FT FT PROPEP 19 30 Activation peptide.
 FT FT CHAIN 31 275 Trypsase beta-2.
 FT FT ACT_SITE 74 74 Charge relay system.
 FT FT ACT_SITE 121 121 Charge relay system.
 FT FT ACT_SITE 224 224 Charge relay system.
 FT FT DISULFID 59 75
 FT FT DISULFID 155 230
 FT FT DISULFID 188 211
 FT FT DISULFID 220 248
 FT FT CARBOHYD 233 233
 FT FT VARIANT 51 53
 FT FT CONFLICT 132 132 N-linked (GlcNAc...) (Potential).
 FT FT STRAND 32 32 HGP -> RDR (in beta-III).
 FT FT STRAND 35 36 /FPIID=VAR 012104.
 FT FT TURN 39 40
 FT FT TURN 40 40
 FT FT TURN 43 44
 FT FT STRAND 45 51
 FT FT STRAND 55 65
 FT FT TURN 66 67
 FT FT STRAND 68 71
 FT FT HELIX 73 75
 FT FT TURN 76 76
 FT FT STRAND 79 79
 FT FT HELIX 83 85
 FT FT STRAND 86 90
 FT FT STRAND 94 94
 FT FT TURN 95 98
 FT FT STRAND 102 102
 FT FT STRAND 104 109
 FT FT TURN 111 112
 FT FT HELIX 116 118
 FT FT STRAND 123 127

FT STRAND 141 141
 FT TURN 145 146
 FT TURN 151 152
 FT STRAND 155 159
 FT STRAND 164 164
 FT TURN 165 166
 FT STRAND 167 167
 FT TURN 171 172
 FT STRAND 174 174
 FT STRAND 176 180
 FT STRAND 182 183
 FT HELIX 185 191
 FT TURN 192 195
 FT STRAND 196 196
 FT TURN 200 201
 FT TURN 207 208
 FT STRAND 209 212
 FT STRAND 214 214
 FT TURN 215 216
 FT STRAND 217 218
 FT TURN 221 222
 FT TURN 224 225
 FT STRAND 227 232
 FT TURN 233 234
 FT STRAND 235 244
 FT STRAND 250 250
 FT TURN 251 252
 FT STRAND 253 253
 FT STRAND 255 259
 FT HELIX 260 263
 FT HELIX 264 270
 SQ SEQUENCE 275 AA; 30529 MW; 2857396C51F5C7A0 CRC64;

Query Match 92.9%; Score 1405; DB 1; Length 275;
 Best Local Similarity 92.0%; Pred. No. 7.7e-114;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Oy 1 MESLILLALPVLASRAYAPAPVQALQAGIVGQGEADRSKMPQVSLRVRDRYMMHFCG 60
 1 MNLILLLPLVLSASRAYAAPAGQMLQRFVGIIVGQGEADRSKMPQVSLRVRDRYMMHFCG 60
 61 GSLHPQWVLTAAHCGVDPVXDLAALRVQLRBDHLHYDDQLLPSRIIVHPQFYTAQLGA 120
 61 GSLHPQWVLTAAHCGVDPVXDLAALRVQLRBDHLHYDDQLLPSRIIVHPQFYTAQLGA 120
 Oy 121 DIALLLEBPVNVSSRVHTVMLPPASFTFPFGMPCWVTGKMGDNDNDEBLLPPFPPLKQYKV 180
 121 DIALLLEBPVNVSSRVHTVMLPPASFTFPFGMPCWVTGKMGDNDNDEBLLPPFPPLKQYKV 180
 Oy 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCKGDSSGGLVCKVNGTWLQAG 240
 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCKGDSSGGLVCKVNGTWLQAG 240
 Db 241 VVSWDEGCAQPRRPGITRYRVTYLDMIHVVYPKP 275
 241 VVSWDEGCAQPRRPGITRYRVTYLDMIHVVYPKP 275
 Db 241 VVSWGEGCAQPRRPGITRYRVTYLDMIHVVYPKP 275
 241 VVSWGEGCAQPRRPGITRYRVTYLDMIHVVYPKP 275

RESULR 7
 O6NZY1
 ID O6NZY1 PRELIMINARY; PRT; 282 AA.
 AC O6NZY1;
 DR 05-JUL-2004 (TREMBlrel. 27, Created)
 DR 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DR 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE TP5B2 protein.
 GN Name=TP5B2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX MEDLINE=22388957; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshull S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina B., Farmer A., Rubin G.M., Hong L., Stappelen M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hultky S.W., Villalón D.K., Murray D.M., Sodergren B.J., Lu X., Gibbs R.A., Wahby J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Gimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RT [2]

RL SEQUENCE FROM N.A.

RP TISSUE=Brain;

RC Director MGC Project;

RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to peptidase family S1.

DR EMBL; BC065923; AAH65923.1; -

DR HSSP; P20231; 1AOL.

DR GO; GO:004263; F:chymotrypsin activity; IEA.

DR GO; GO:0004295; F:protease activity; IEA.

DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR009003; Pept_Ser_Cys.

DR Pfam; PF00089; Trypsin; 1_

DR SMART; SMART; CHYMOTRYPSIN.

DR SMART; SM00020; TRYP_Spec; 1.

DR PROSITE; PS50240; TRYP_SIN_DOM; 1.

DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYP_SIN_SER; 1.

KW Hydrolyase; Protease; Serine protease.

KM SEQUENCE 282 AA; 31296 MW; 9727DC52798986CA CRC64;

SO

Query Match 92.9%; Score 1405; DB 2; Length 282;

Best Local Similarity 92.0%; Pred. No. 7.9e-114;

Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60

DB 8 MNLNLLALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 67

QY 61 GSLIHPOVWVLTAAHCGPVDKDLATLRVQLREOHLYYQDQLLPSRSRIIVHPQFYIQTGA 120

DB 68 GSLIHPOVWVLTAAHCGPVDKDLAALRVQDLREOHLYYQDQLLPSRSRIIVHPQFYIQTGA 127

QY 121 DIALLELEBPVNISSRVTHTVMLPPASETFPPGMPQVLTGWGDVNDDELPPFPFLKQYKV 180

DB 128 DIALLELEBPVNISSRVTHTVMLPPASETFPPGMPQVLTGWGDVNDDELPPFPFLKQYKV 187

QY 181 PIMENHICDAKXNHLGAYTGDVRIIRDDMLCAQNSQRDSCKDGSGPLVCKVNGTWLQAG 240

DB 188 PIMENHICDAKXNHLGAYTGDVRIIRDDMLCAQNSQRDSCKDGSGPLVCKVNGTWLQAG 247

QY 241 VVSWDEGCAQPNRPGIYTRVTVYLDWVHNYVPKKP 275

DB 248 VVSWDEGCAQPNRPGIYTRVTVYLDWVHNYVPKKP 282

RESULT 8

Q6FHB8 PRELIMINARY; PRT; 275 AA.

AC Q6FHB8;

DT 05-JUN-2004 (TREMBLrel. 27, Created)

DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update)

DE TP5B2 protein (Fragment).

GN Name=TP5B2;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RA Halleck A., Ebert L., Mkundinya M., Schick M., Eisenstein S., Kernbert P., Katzung K., Scharten R., Shen B., Henze S., Mar W., Korn B., Zhu D., Hu Y., Labaer J.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to peptidase family S1.

DR EMBL; CR541836; CAG46635.1; -

DR GO; GO:004263; F:chymotrypsin activity; IEA.

DR GO; GO:0004295; F:protease activity; IEA.

DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001254; Peptidase_S1.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR009003; Pept_Ser_Cys.

DR Pfam; PF00089; Trypsin; 1_

DR SMART; SMART; CHYMOTRYPSIN.

DR SMART; SM00020; TRYP_Spec; 1.

DR PROSITE; PS50240; TRYP_SIN_DOM; 1.

DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYP_SIN_SER; 1.

KW Hydrolyase; Protease; Serine protease.

FT NON_TER 275 AA; 30517 MW; ABD124C42938688BD CRC64;

SO

Query Match 92.8%; Score 1403; DB 2; Length 275;

Best Local Similarity 91.6%; Pred. No. 1.2e-113;

Matches 252; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60

DB 1 MNLNLLALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60

QY 61 GSLIHPOVWVLTAAHCGPVDKDLATLRVQLREOHLYYQDQLLPSRSRIIVHPQFYIQTGA 120

DB 61 GSLIHPOVWVLTAAHCGPVDKDLAALRVQDLREOHLYYQDQLLPSRSRIIVHPQFYIQTGA 120

QY 121 DIALLELEBPVNISSRVTHTVMLPPASETFPPGMPQVLTGWGDVNDDELPPFPFLKQYKV 180

DB 121 DIALLELEBPVNISSRVTHTVMLPPASETFPPGMPQVLTGWGDVNDDELPPFPFLKQYKV 180

QY 181 PIMENHICDAKXNHLGAYTGDVRIIRDDMLCAQNSQRDSCKDGSGPLVCKVNGTWLQAG 240

DB 181 PIMENHICDAKXNHLGAYTGDVRIIRDDMLCAQNSQRDSCKDGSGPLVCKVNGTWLQAG 240

QY 241 VVSWDEGCAQPNRPGIYTRVTVYLDWVHNYVPKKP 275

DB 241 VVSWDEGCAQPNRPGIYTRVTVYLDWVHNYVPKKP 275

RESULT 9

Q6FRZ6 PRELIMINARY; PRT; 275 AA.

AC Q6FRZ6;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Trypsase 1.

GN Name=tryptasec;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

CC response of this cell type.
 CC -1- CATABOLIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
 CC with more restricted specificity than trypsin.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=The alternative splicing event is due to a G to A point
 CC mutation at the exon 2/intron 2 splice site and causes loss of
 CC protein expression. The alternatively spliced transcript is only
 CC found in C57BL/6 mouse;
 CC Name=1;
 CC IsoId=Q02844-1; Sequence=Displayed;
 CC Name=2; Synonyms=truncated;
 CC IsoId=Q02844-2; Sequence=VSP_005378, VSP_005379;
 CC -1- DEVELOPMENTAL STAGE: Is not expressed in mature serosal or mucosal
 CC mast cells and is expressed only transiently at an early stage of
 CC in vitro mast cell differentiation.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Trypsase
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L00654; AAA3993.1; -
 CC EMBL; L00653; AAA3992.1; -
 CC EMBL; U42405; AAA97874.1; -
 CC EMBL; U42406; AAA97875.1; -
 CC PIR; A47246; A47246.
 CC HSSP; P15157; LTTO.
 CC MEROPS; S01.026; -
 CC MGD; MG1:96943; MGPL7.
 CC DR InterPro: IPR009003; Pept Ser Cys.
 CC DR InterPro: IPR001254; Peptidase_S1.
 CC DR InterPro: IPR001314; Peptidase_S1A.
 CC DR Pfam; PFO0089; Trypsin; 1.
 CC DR PRINTS; PR00722; CHYMOTRYPSIN.
 CC DR SMART; SM0020; TRYP_SPC; 1.
 CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE; PS00135; TRYPSIN_SRR; 1.
 CC KW Alternative splicing; Glycoprotein; Hydrolyase; Serine protease;
 CC Signal; Zymogen.
 CC FT SIGNAL 1 18 Potential.
 CC FT PROPEP 19 28 Activation peptide (Potential).
 CC FT CHAIN 29 273 Mast cell protease 7.
 CC FT ACT_SITE 72 72 Charge relay system (By similarity).
 CC FT ACT_SITE 119 119 Charge relay system (By similarity).
 CC FT ACT_SITE 222 222 Charge relay system (By similarity).
 CC FT DISULFID 57 73 By similarity.
 CC FT DISULFID 153 228 By similarity.
 CC FT DISULFID 186 209 By similarity.
 CC FT DISULFID 218 246 By similarity.
 CC FT CARBOHYD 49 49 N-linked (GlcNAc...) (Potential).
 CC FT CARBOHYD 130 130 N-linked (GlcNAc...) (Potential).
 CC FT VARSPLIC 44 46 VSL -> GCC (in isoform 2).
 CC FT VARSPLIC 47 273 Missing (in isoform 2).
 CC FT SEQUENCE 273 AA; 30337 MW; 50BCBA4755294205E CRC64;
 CC Query Match 75.0%; Score 1134; DB 1; Length 273;
 CC Best Local Similarity 74.4%; Pred. No. 2, 7e-90;
 CC Matches 203; Conservative 23; Mismatches 45; Indels 2; Gaps 1;

DB 59 GSLIHPQWVLTAAHCIGPVDKDLATLRVQLREQHLYYQDQLLPSVRIIVHPQFIITGGA 118
 QY 121 DIALLLEFEPVNISSRHTTWLPPASSTFPPGMCWVTGMGDVNDDELPPLPKQKYV 180
 DB 119 DIALKLTNPVNIIDYVHPPLPPASSTFPPGTLCTVYGMWINDVNLPPPLKVEQV 178
 QY 181 PIMENHICDAVYHGAATYGDVRIIRPDMLCAGNSQSDSCKSGSGPLVCKVNGTMLQAG 240
 DB 179 PIENHLDLTKYHGLITGDVNIIVRDMLCAGNBEHSDSCGDSGSPLVCKVNBETMLQAG 238
 QY 241 VVSWDEGCAQPNRPFGIYTRVTVYLDWTHHVVPK 273
 DB 239 VVSWGEGCAQPNRPFGIYTRVTVYLDWTHHVVPK 271
 RESULT 12
 MCT6_MOUSE STANDARD; PRT; 276 AA.
 AC P21845; 061962;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Mast cell protease 6 precursor (BC 3.4.21.59) (MCP-6) (Trypsase).
 GN Name=MCP6;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139682; PubMed=1995638;
 RA Reynolds D.S., Gurley D.S., Austen K.F., Serafin W.E.;
 RT "Cloning of the cDNA and gene of mouse mast cell protease-6.
 RT Transcription by progenitor mast cells and mast cells of the
 RT connective tissue subclass.";
 RL J. Biol. Chem. 266:3847-3853(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Leaden X A1;
 RX MEDLINE=94023807; PubMed=8210998;
 RA Huang R., Abdrink M., Gobl A.E., Nilsson G., Aveskog M., Larsson L.G.,
 RA Nilsson K., Hellman L.;
 RT "Expression of a mast cell tryptase in the human monocytic cell lines
 RT U-937 and Mono Mac 6.";
 RL Scand. J. Immunol. 38:359-367(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=Leaden X A1;
 RX MEDLINE=95048582; PubMed=7959952;
 RA Huang R., Hellman L.T.;
 RT "Genes for mast-cell serine protease and their molecular evolution.";
 RL Immunogenetics 40:397-414(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC STRAIN=FVB/N. TTSUB=COLON.
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.D., Collins F.S., Wagner G.J., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tomihata S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

QY 1 MSLIILALPVAASRKYARPAVQVLAQOAGTYGCGEAPASRKKPQVSLAVRPRYMMHRCG 60
 DB 1 MTKLILTLPLRSLISLVHARBP-AWTRRGIVGCGEBAHKNKWPQVSLANDTYMMHRCG 58
 QY 61 GSLIHPQWVLTAAHCIGPVDKDLATLRVQLREQHLYYQDQLLPSVRIIVHPQFIITGGA 120

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [5]
 RP SEQUENCE OF 32-54.
 RX MEDLINE=90222202; PubMed=2326280;
 RA Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,
 RA Seratin W.E.;
 RT "Different mouse mast cell populations express various combinations of
 at least six distinct mast cell serine proteases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990).
 CC -1- FUNCTION: Trypsin is the major neutral protease present in mast
 cells and is secreted upon the coupled activation-degranulation
 response of this cell type.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|- , Lys-|- , but
 with more restricted specificity than trypsin.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P21845-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P21845-2; Sequence=VSP_005376, VSP_005377;
 CC Note=Probably non functional.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Trypsin
 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M57626; AAA39988.1; -;
 DR EMBL; M57625; AAA39987.1; -;
 DR EMBL; L31853; AAA39725.1; -;
 DR EMBL; X78542; CA55288.1; -;
 DR EMBL; BC024374; AAH24374.1; -;
 DR PIR; A38654; A38654.
 DR PIR; I48685; I48685.
 DR HSSP; P20231; IAO1.
 DR MEROPS; S01.025; --.
 DR WGI; M61.96942; M61.96942.
 DR GO; GO:0008201; F:heparin binding; IDA.
 DR GO; GO:0005115; F:protein binding; IPI.
 DR GO; GO:0030019; F:trypsin activity; IDA.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00134; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00135; TRYP_SIN_HIS; 1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 KW Alternative splicing; Direct protein sequencing; Glycoprotein;
 KW Hydrolyase; Serine protease; Signal; Zymogen.
 FT SIGNAL 1 21
 FT PROPEP 22 21
 FT CHAIN 32 276
 FT ACT_SITE 75 75 Mast cell protease 6.
 FT ACT_SITE 122 122 Charge relay system (By similarity).
 FT ACT_SITE 225 225 Charge relay system (By similarity).
 FT DISUFID 60 76 Charge relay system (By similarity).
 FT DISUFID 156 231 By similarity.
 FT DISUFID 189 212 By similarity.
 FT DISUFID 221 249 By similarity.
 FT CARBOHYD 133 133 N-linked (GlcNAc...) (Potential).
 FT VARSPDIC 223 230 GDSGGPLV -> PFCIGDDI (in isoform Short).
 FT VARSPDIC 231 276 /FTID=VSP_005376.
 FT VARSPDIC 231 276 Missing (in isoform Short).

FT SEQUENCE 276 AA; 30927 MW; 525B2C9A04A17200 CRC64;
 SQ 74.9%; Score 1133; DB 1; Length 276;
 Query Match Best Local Similarity 73.3%; Pred. No. 3.3e-90;
 Matches 198; Conservative 28; Mismatches 42; Indels 2; Gaps 1;
 QY 4 ILLALPLVIASRAVAAPVQALQAGIVGGQAEPRSRMPCWQSLRVRDRYMMHFQGGSL 63
 DB 7 ILLMMLSLIASLIVSAPRPP--ANCRVGIIVGHEASESEKMPQVSLRFLKLNWIHFQGGSL 64
 QY 64 IHPQVTLTAHQLGQVDVLDLTLRYQLREOHLYYDQLLPSRIIVHPOFIYIQGADIA 123
 DB 65 IHPQVTLTAHCVGPHINSPLQFRVQLREQYLYGDDQLSLRIVVHHPHYTTAEGGADVA 124
 QY 124 ILELEPWNISRRVHTVMPPASETFPPGMPCTVWGMDVNDDEPLPPFPKQKVPIM 183
 DB 125 ILELEVPVNVSHIHPISLPPASERFPFGTSCWVYGMWDINDDEPLPPFPKQKVPYIV 184
 QY 184 ENHICDAPKHHGAYTGDDVRIIRDMLCAGNSQRDSCKGGSGPLVCKVNGTMLQAGVVS 243
 DB 185 ENSLCDRKRKHGTGLYGDPPFIVHDGMLCAGNTRDSCCGDSGGLVCKVNGTMLQAGVVS 244
 QY 244 WDEGCAQPNRPGIYRVTYTLDMTHHYPK 273
 DB 245 WDEGCAQPNRPGIYRVTYTLDMTHHYPK 274
 RESULT 13
 TRYT CANFA
 ID TRYT CANFA STANDARD; PRT; 275 AA.
 AC P15944;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Trypsin precursor (BC 3.4.21.59).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 ON NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89352460; PubMed=2504277;
 RA Vanderslice P., Craik C.S., Madel J.A., Caughey G.H.;
 RT "Molecular cloning of dog mast cell trypsin and a related protease:
 structural evidence of a unique mode of serine protease activation.";
 RL Biochemistry 28:4148-4155(1989).
 CC -1- FUNCTION: Trypsin is the major neutral protease present in mast
 cells and is secreted upon the coupled activation-degranulation
 response of this cell type.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|- , Lys-|- , but
 with more restricted specificity than trypsin.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Released from the secretory granules upon
 mast cell activation.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family. Trypsin
 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M24664; AAA30854.1; -;
 DR PIR; A32410; A32410.
 DR HSSP; P20231; IAO1.
 DR MEROPS; S01.118; --.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.

DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART; SM0020; TRYP_Spc; 1.
 DR PROSITE; PSS0240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 KW Glycoprotein; Hydrolase; Serine protease; signal; zymogen.
 FT SIGNAL 1 20 Potential. (By similarity).
 FT PROPEP 1 30 Activation peptide (By similarity).
 FT CHAIN 31 275 Trypsase.
 FT ACT_SITE 74 74 Charge relay system (By similarity).
 FT ACT_SITE 121 121 Charge relay system (By similarity).
 FT ACT_SITE 224 224 Charge relay system (By similarity).
 FT DISULFID 59 75 By similarity.
 FT DISULFID 155 230 By similarity.
 FT DISULFID 188 211 By similarity.
 FT DISULFID 220 248 By similarity.
 FT CAROHYD 132 132 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 275 AA; 30088 MW; C38869251F248D5B CRC64;
 Query Match 74.8%; Score 1131; DB 1; Length 275;
 Best Local Similarity 72.7%; Pred. No. 4,9e-90;
 Matches 200; Conservative 30; Mismatches 45; Indels 0; Gaps 0;

QY 1 MSLLLLALPVLASRAVAAPVQALQAGIVGGQEARPSKMPQVSRVDRRMMHFCG 60
 DB 1 MSPPLVALALAGSLVPSVPAFGQALQKRVIGVGRAPSSKMPQVSLKQGYMWHICG 60
 QY 61 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 120
 DB 61 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 120
 QY 61 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 120
 DB 61 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 120
 QY 121 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 180
 DB 121 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 180
 QY 121 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 180
 DB 121 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 180
 QY 181 PIMENHICDAKYHIGAYTGDDVRIIRDDMLCAGNSQSDSCGDSGGLVCKVNGTWLQAG 240
 DB 181 PIMENHICDAKYHIGAYTGDDVRIIRDDMLCAGNSQSDSCGDSGGLVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMVHRYVPK 275
 DB 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMVHRYVPK 275
 RESULT 14
 Q921N4 PRELIMINARY; PRT; 273 AA.
 ID Q921N4
 AC Q921N4 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Mast cell protease 7.
 GN Name=Mcp7;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ussidri T.B., Toshlyuk S., Carninci P., Franke C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abrahamsen R.D., Mullighy S.J.,
 RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richarde S., Worley K.C., Hale A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettleman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myere R.M., Butcherfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RA and mouse cDNA sequences".
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CZECH II; TISSUE=Mammary tumor;
 RC Strausberg R.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL - SIMILARITY: Belongs to peptidase family 11.
 CC EMBL; BC011328; AAH11328.1; -
 DR HSSP; P15157; ILVO.
 DR MGD; MGI:96943; Mcp7.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM0020; TRYP_Spc; 1.
 DR PROSITE; PSS0240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 273 AA; 30332 MW; 514174765140358B CRC64;
 Query Match 74.3%; Score 1123; DB 2; Length 273;
 Best Local Similarity 70.4%; Pred. No. 2,4e-89;
 Matches 202; Conservative 22; Mismatches 47; Indels 2; Gaps 1;

QY 1 MSLLLLALPVLASRAVAAPVQALQAGIVGGQEARPSKMPQVSRVDRRMMHFCG 60
 DB 1 MKKLLLLPLPLSSVYLAAPG--AMTREGIVGGQEARGNKMPQVSRVDRRMMHFCG 58
 QY 61 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 120
 DB 59 GSLIHPQWVLTAAHCLGPPVDKDLATLRVQLREOHLYYQDQLPVSRIIVHPQFYIITG 118
 QY 121 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 180
 DB 119 DIALLELEBPVNISSRVHTVMLPASSETPPGMPGCVTGWGVNDDEPLPPPPPLKQYV 178
 QY 181 PIMENHICDAKYHIGAYTGDDVRIIRDDMLCAGNSQSDSCGDSGGLVCKVNGTWLQAG 240
 DB 179 PIMENHICDAKYHIGAYTGDDVRIIRDDMLCAGNSQSDSCGDSGGLVCKVNGTWLQAG 238
 QY 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMVHRYVPK 273
 DB 239 VVSWDEGCAQPNRPGIYTRVTVYLLDMVHRYVPK 271
 RESULT 15
 Q96RZ7 PRELIMINARY; PRT; 233 AA.
 ID Q96RZ7
 AC Q96RZ7 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Mast cell tryptase beta III.
 GN Name=tryptaseb;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21096910; PubMed=1157797; DOI=10.1093/hmg/10.4.339;
 RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
 RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
 RA Higgs D.R.;
 RT "Sequence, structure and pathology of the fully annotated terminal 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 23:39:06 ; Search time 615 Seconds
(without alignments)
2925.772 Million cell updates/sec

Title: US-09-598-982C-52

Sequence: 1 MSLTLLALPVLASAYAAP.....YTRVYVYLDIMHRYVPPKP 275

Scoring table:
BLOSSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=x1p
-CGN2_1/USPTO.spool.p/US09598982/rnmpat.25082005.165827.1602/app.query.faeta_1.455
-DB=Published Applications NA -OPMT=faetrap -SUPFIX=trmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -STRAPT=1 -END=-1 -MATRIX=bloms62
-FRANS=human40.ccd1 -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USR=US09598982 @CGN 1.1 723 @rnmpat.25082005.165827.1602
-NCPUS=6 -ICPU=3 -NO NMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Match, Length, DB, ID, Description. Contains 45 entries of search results.

ALIGNMENTS

RESULT 1
US-10-352-684A-45
; Sequence 45, Application US/10352684A
; Publication No. US20030215552A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Healy, Aileen
; APPLICANT: Welch, Nadine S.
; APPLICANT: Kelly, Louise M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12203, 13906,
; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
; TITLE OF INVENTION: 19811, 16816, 270, 1410, 137, 1871, 13051, 1847, 1849,
; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
; FILE REFERENCE: MP102-019P1RNONMIM

Pred. No. is the number of results predicted by chance to have a

Db 199 GGCCTCCATCCACCCCAAGGGGGTGGACCCGGGGGGCACTGGCTGGGACCGGACCGTC
 Qy 81 LysApPLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrThrGlnAspGln 100
 Db 258 AAGGATCTGGCCACCTCCAGGGGTGCAACTGCCGGAGCAGCACCTCTACTACCAAGACCG 317
 Qy 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 Db 318 CTGCTGCCAGTACGACGATCTCTGTGACCCACAGTTTCAATCAATCCAGACTGGAGCG 377
 Qy 121 AspIleAlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrVal 140
 Db 378 GATATCGCCCTCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGGGTCCACACCGGTC 437
 Qy 141 MetLeuProProAlaSerGlnUthrPheProProGlyMetProCysTyrValIthrGlyTyr 160
 Db 438 ATGCTGCCCCCTCCCGGAGACCTTCCCGGGGGATGCCGTGGTCACTGGCTGG 497
 Qy 161 G1yAspValAspAsnAspGlnProLeuProProProPheProLeuIlyGlnValIlyVal 180
 Db 498 GCGGATGAGCAATGATAGCCCTCCACCGCATTTCCCTGAAGCAGGTGAAGGTC 557
 Qy 181 ProIleMetGluAsnHisIleCysAspAlaIlyTyrHisLeuGlyAlaTyrThrGlyAsp 200
 Db 558 CCCATTAATGAAAACCACTTTGTGACGGAATAACACCTTGGCCCTACACCGGGAGAC 617
 Qy 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 Db 618 GACGTCGGCATCATCCGTGACGATGCTGTGTGCCGGGAAACAGCCAGGAGCACTCCGTC 677
 Qy 221 IyG1yAspSerGlyGlyProLeuValCysIlyValAsnGlyIlyThrIleuGlnIlyGly 240
 Db 678 AAGGGCGACTCTGGAGGGCCCTGGTGTGCAAGTGAATGGCACTGGCTACAGCGGGGC 737
 Qy 241 ValValSerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIlyTyrThrArgVal 260
 Db 738 GTGTCAGCTGGAGCGAGGGCTGTGCCAGCCCAACCGGCCCTGGCATCTACACCGGTGTC 797
 Qy 261 ThrTyrTyrLeuAspTrpIleHisIlyValProIlyIlyPro 275
 Db 798 ACCTACTACTGAGCTGGATCCACCACTATGTCTCCCAAAAAGCCG 842

RESULT 3
 US-10-956-157-1042
 / Sequence 1042, Application US/10956157
 / Publication No. US20050118625A1
 / GENERAL INFORMATION:
 / APPLICANT: Wyeth
 / APPLICANT: Mounts, William
 / TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 / TITLE OF INVENTION: HUMAN OSTEOBLASTS AND HUMAN PROTEASES
 / FILE REFERENCE: 031896-04300 (AM 101081)
 / CURRENT APPLICATION NUMBER: US/10/956, 157
 / CURRENT FILING DATE: 2004-10-04
 / NUMBER OF SEQ ID NOS: 319805
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 1042
 / LENGTH: 1158
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-10-956-157-1042

Alignment Scores:
 Pred. No.: 5,54e-171 Length: 1158
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-598-982c-52 (1-275) x US-10-956-157-1042 (1-1158)

Qy 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTyrAlaIlePro 20
 Db 18 ATGCTGAGCCCTCGCTGCTGGCCCTGTCCGCTCTGGGGAGACCCGCGCTTACAGCGGCCCT 77
 Qy 21 AlaProValGlnAlaLeuGlnGlnIlyIleValGlyGlyGlnGlnAlaProArgSer 40
 Db 78 GCCCCAGTCCAGGCCCTCCAGCAAGCCGGGTATCCCTCCGGGGGTCCAGGAGGCCCCCAAGGAGC 137
 Qy 41 IystrProTrpGlnValSerLeuArgValIleArgAspArgTyrTrpMetHisPheCysGly 60
 Db 138 AAGTGGCCCTGGCAGGTGAGGCTGAGAGTCCGCAAGATCTGATCAATCTGATCTTCTGGGG 197
 Qy 61 GlySerLeuIleHisProGlnTrpValIleuThrAlaAlaHisCysLeuGlyProAspVal 80
 Db 198 GGCCTCCCTCATCCACCCCACTGGGTGTGACCCGGGGGCACTGCTGGAGCCGGAGCGTC 257
 Qy 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGln 100
 Db 258 AAGGATCTGGCCACCTCCAGGGGTCAACTGCGGGAGGACACTCTACTACCAAGACCGAG 317
 Qy 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 Db 318 CTGCTGCCAGTACGACGATCTCTGTGACCCACAGTTTCAATCAATCCAGACTGGAGCG 377
 Qy 121 AspIleAlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrVal 140
 Db 378 GATATCGCCCTCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGGGTCCACACCGGTC 437
 Qy 141 MetLeuProProAlaSerGlnUthrPheProProGlyMetProCysTyrValIthrGlyTyr 160
 Db 438 ATGCTGCCCCCTCCCGGAGACCTTCCCGGGGGATGCCGTGGTCACTGGCTGG 497
 Qy 161 G1yAspValAspAsnAspGlnProLeuProProProPheProLeuIlyGlnValIlyVal 180
 Db 498 GCGGATGAGCAATGATAGCCCTCCACCGCATTTCCCTGAAGCAGGTGAAGGTC 557
 Qy 181 ProIleMetGluAsnHisIleCysAspAlaIlyTyrHisLeuGlyAlaTyrThrGlyAsp 200
 Db 558 CCCATTAATGAAAACCACTTTGTGACGGAATAACACCTTGGCCCTACACCGGGAGAC 617
 Qy 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 Db 618 GACGTCGGCATCATCCGTGACGATGCTGTGTGCCGGGAAACAGCCAGGAGCACTCCGTC 677
 Qy 221 IyG1yAspSerGlyGlyProLeuValCysIlyValAsnGlyIlyThrIleuGlnIlyGly 240
 Db 678 AAGGGCGACTCTGGAGGGCCCTGGTGTGCAAGTGAATGGCACTGGCTACAGCGGGGC 737
 Qy 241 ValValSerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIlyTyrThrArgVal 260
 Db 738 GTGTCAGCTGGAGCGAGGGCTGTGCCAGCCCAACCGGCCCTGGCATCTACACCGGTGTC 797
 Qy 261 ThrTyrTyrLeuAspTrpIleHisIlyValProIlyIlyPro 275
 Db 798 ACCTACTACTGAGCTGGATCCACCACTATGTCTCCCAAAAAGCCG 842

RESULT 4
 US-10-723-860-6799
 / Sequence 6799, Application US/10723860
 / Publication No. US20040253606A1
 / GENERAL INFORMATION:
 / APPLICANT: Aziz, Natsaba
 / APPLICANT: Ghisburg, Wendy M.
 / TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
 / TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
 / FILE REFERENCE: 05882.0193.NUS01
 / CURRENT APPLICATION NUMBER: US/10/723, 860
 / CURRENT FILING DATE: 2003-11-26
 / PRIORITY APPLICATION NUMBER: 60/429, 739
 / NUMBER OF SEQ ID NOS: 8393
 / SOFTWARE: PatentIn version 3.2

Alignment Scores:
 Pred. No.: 5,54e-171 Length: 1158
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-598-982c-52 (1-275) x US-10-956-157-1042 (1-1158)

NAME/KEY: CDS
LOCATION: (1)..(825)
US-10-287-226-89

Alignment Scores:
Pred. No.: 2,05e-159 Length: 828
Score: 1414.00 Matches: 255
Percent Similarity: 95.64% Conservat: 8
Best Local Similarity: 92.73% Mismatches: 12
Query Match: 93.52% Indels: 0
DB: 18 Gaps: 0

US-09-598-982c-52 (1-275) x US-10-287-226-89 (1-828)

1 MetLeuSerLeuLeuLeuLeuLeuLeuProValIleSerArgAlaThrAlaIlePro 20
1 ARGCTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 60
21 AlaProValAlaAlaLeuGlnGlnAlaGlyIleValAlaGlyGlnGlnAlaProArgSer 40
61 GCCCCAGGCGCCAGGCGCCCTGAGCAAGGGCCATTGTTGGGGGCGAGAGGCCCCCGAGAGC 120
41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTrpTrpMetHisSphCysGly 60
121 AAGTGGCCCTGGCAGGTGAGCTGAGAGAGTCCGGCGCCCACTACTGAGATGCACTTCTGGGG 180
61 GlySerLeuIleHisProGlnTrpValIleThrAlaAlaHisCysLeuGlyProAspVal 80
181 GGCTCCCTATCCACCCCGAGGGGTGCTAACCGCGCCGCTGCGAGAACCGGAGCAGTCC 240
81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisIleuTrpTrpGlnAspGln 100
241 AAGGATCTGGCCCGCTCAGGGTGGCACTGGGGAGCAGCACTTACTACAGAGCCAG 300
101 LeuLeuProValSerArgIleIleValHisProGlnPheTrpIleIleGlnThrGlyAla 120
301 CTGGCCCGGTGAGCAGATCCGTGCAACCCAGTCTTACTACATGATCCAGACCGGGGGG 360
121 AspIleAlaLeuLeuGlnLeuGlnGluProValAsnIleSerSerArgValHisThrVal 140
361 GACATCGCCCTCTGAGGCTGGAGGGAGCCCGAGACATCTCCAGCAGCAGCAGCGGTC 420
141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrpValThrGlyTrp 160
421 ACGCTGCCCCCTGCTCGGAGACCTTCCCGGGGAGTCCCGTGGTCACTGGCTGG 480
161 GlyAspValAspAsnAspGlnProLeuProProProPheProLeuIleValIleVal 180
481 GCGGACGTGGACAATATAGAGGGCCCTCCAGCCGCACTTCTGAGAGGAGTGAAGGTC 540
181 ProIleMetGlnAsnHisIleCysAspAlaLysTrpHisLeuGlyAlaTrpThrGlyAsp 200
541 CCCAATATGAAACCAACCACTTTGTGACGCAAAATATCCACTTGGCGCCCTCAACCGGAGAGC 600
201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnAsnArgAspSerCys 220
601 GACGTCGCCGATGTCTCGTGAACATGCTGTGTGCGGGAACAACCGGGGAGACTCATGTC 660
221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
661 CAGGGGGAATCCGGAGGGGCGCCCGTGGTGCAGAGTGAATGGGACCTGGGCTGAGAGGGGGC 720
241 ValValSerTrpAspGlnGlyCysAlaGlnProAsnAspProGlyIleTrpThrArgVal 260
721 GTGGTCAGCTGGGGCGAGGCTGTGCTGCCAGCCCAACCGGCGCTGGCATCTACACCCCGTGTTC 780
261 ThrTrpTrpLeuAspTrpIleHisHisTrpValProLysLysPro 275
781 ACCTACTACTGGAGCTGATCCACACTATGTCCCAAAAAGCCG 825

Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Bergin, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chandhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eisen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Esha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khrantsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Maljankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Miller, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Murajidhara,
APPLICANT: Patburajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Verneil, Corine A.M.,
APPLICANT: Zerhusen, Bryan D.,
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: Curseqdist version 0.1
SEQ ID NO 91
LENGTH: 828
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(825)

RESULT 6
US-10-287-226-91
Sequence 91, Application US/10287226

QY 241 ValValSerTPPaspGluGlyCysAlaGlnProAsnArgProGlyLeuThrArgVal 260
 DB 773 GTGGTCAAGCTGGGGCGAGGGCTGTGTCCAGCCCAACCGGCTGGCACTTACACCCCGTGTTC 832
 QY 261 ThrTYTYLeuAspTRPILeHisHisTYValProLYSLYsPro 275
 DB 893 ACCTACTACTTGGACTGGATCCACCACTATGTCTCCCAAAAAGCCG 877

RESULT 8

US-10-352-684A-47
 ; Sequence 47, Application US/10352684A
 ; Publication No. US20030215452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals Inc.
 ; APPLICANT: Carroll, Joseph W.
 ; APPLICANT: Healy, Aileen
 ; APPLICANT: Welch, Nadine S.
 ; APPLICANT: Kelly, Louise M.
 ; APPLICANT: Louis, M.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
 ; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
 ; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
 ; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
 ; FILE REFERENCE: MPI02-019P1RNONMIM
 ; CURRENT APPLICATION NUMBER: US/10/352,684A
 ; CURRENT FILING DATE: 2003-01-28
 ; PRIOR APPLICATION NUMBER: US 60/354,333
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: US 60/360,258
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/364,476
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: US 60/375,626
 ; PRIOR FILING DATE: 2002-04-26
 ; PRIOR APPLICATION NUMBER: US 60/386,494
 ; PRIOR FILING DATE: 2002-06-06
 ; PRIOR APPLICATION NUMBER: US 60/390,965
 ; PRIOR FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: US 60/392,480
 ; PRIOR FILING DATE: 2002-06-28
 ; PRIOR APPLICATION NUMBER: US 60/394,128
 ; PRIOR FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: US 60/399,783
 ; PRIOR FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: US 60/403,221
 ; PRIOR FILING DATE: 2002-08-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 1143
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (6)...(833)
 ; US-10-352-684A-47

Alignment Scores:

Pred. No.: 3,79e-158 Length: 1143
 Score: 1405.00 Matches: 253
 Percent Similarity: 94.918 Conservative: 8
 Best Local Similarity: 92.008 Mismatches: 14
 Query Match: 92.928 Indels: 0
 DB: 17 Gaps: 0
 US-09-598-982C-52 (1-275) x US-10-352-684A-47 (1-1143)
 QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTYRAlaAlaPro 20
 DB 6 ATGCGTAATCTGCTGTGGCGCTGCCCTTCTTGGCGAGCCGCCCTTACGGGCGCCCT 65
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40

DB 66 GCCCAGGCCAGGCGCTGCAGCCAGGAGTGGATCGTTGGGGGTTCAGAGGCCCCAGAGGC 125
 QY 41 LysTRPProTYRglnValSerLeuArgValArgAspArgTYRTRPMeHisPheCysGly 60
 DB 126 AAGTGGCCCTGGCAGGAGGCTGAGAGTCCAGCCCACTCTGTAAGTCCACTTCTGGCCGG 185
 QY 61 GlySerLeuIleHisProGlnInTRValLeuThrAlaHisCysValGlyProAspVal 80
 DB 186 GACTCCCTCATCCACCCCAAGTGGGTGTGACCCAGCCAGCCAGCTGCTGGAGCCGGACGTC 245
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTYRTRGlnAspGln 100
 DB 246 AAGATCTGGCGCCCTCAGGAGTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 305
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTYRLeuIleGlnInTRGlyVal 120
 DB 306 CTGGCTGCCGCTCACAGGATCATGCTGACCCACAGTTCACACCCGCGCATGCGGAGCGG 365

RESULT 9

US-10-287-226-93
 ; Sequence 93, Application US/10287226
 ; Publication No. US20040086875A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agee, Michele L.,
 ; APPLICANT: Alsbrook, John P.,
 ; APPLICANT: Berghs, Constance,
 ; APPLICANT: Boldog, Ferenc,
 ; APPLICANT: Burgess, Catherine E.,
 ; APPLICANT: Chant, John S.,
 ; APPLICANT: Chandhuri, Amitabha,
 ; APPLICANT: DiPippo, Vincent A.,
 ; APPLICANT: Edinger, Shlomit R.,
 ; APPLICANT: Eisen, Andrew,
 ; APPLICANT: Ellerman, Karen,
 ; APPLICANT: Gangolli, Bsha A.,
 ; APPLICANT: Gorznan, Linda,
 ; APPLICANT: Gerlach, Valerie,
 ; APPLICANT: Ji, Weizhen,
 ; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Khramtsov, Nikolai,
 ; DB 261 ThrTYTYLeuAspTRPILeHisHisTYValProLYSLYsPro 275
 ; DB 786 ACCTACTACTTGGACTGGATCCACCACTATGTCTCCCAAAAAGCCG 832

RESULT 9

QY 241 ValValSerTPPaspGluGlyCysAlaGlnProAsnArgProGlyLeuThrArgVal 260
 DB 726 GTGGTCAAGCTGGGGCGAGGGCTGTGTCCAGCCCAACCGGCTGGCACTTACACCCCGTGTTC 785
 QY 261 ThrTYTYLeuAspTRPILeHisHisTYValProLYSLYsPro 275
 DB 786 ACCTACTACTTGGACTGGATCCACCACTATGTCTCCCAAAAAGCCG 832

```

APPLICANT: Li, Li,
APPLICANT: Malankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Patirajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vermet, Corine A.M.,
APPLICANT: Zerhusen, Bryan D.,
APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining prior application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: Cursesqlst version 0.1
SEQ ID NO 93
LENGTH: 1145
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (8)..(832)
US-10-287-226-93

Alignment Scores:
Pred. No.: 3.8e-158 Length: 1148
Score: 1405.00 Matches: 253
Percent Similarity: 94.91% Conservative: 8
Best Local Similarity: 92.00% Mismatches: 14
Query Match: 92.92% Indels: 0
DB: 18 Gaps: 0

US-09-598-982C-52 (1-275) x US-10-287-226-93 (1-1145)
Qy 1 MetLeuSerIleuLeuLeuAlaLeuProValIleuAlaSerArgAlaTyrAlaIlePro 20
Db 8 ATGCTAATTCGTGCTGCTGGCGCCGCTGCGCGCGCGCCCTACCGCGCCCT 67
Qy 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValIGIYGIYGIInGlnIuAlaProArgSer 40
Db 68 GCCCCAGCCAGCCCTGTCGACGAGCGATCGTTGGGGGTCACGAGGCCCCCGAGAGC 127

```

```

Qy 41 LysTrpProTrpGlnValSerIleuArgValArgAspArgTyrTrpMetHisPheCysGly 60
Db 128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGCCAGCCCACTCTGATGACACTTTCGGCGG 187
Qy 61 GlySerIleuIleHisProGlnTrpValIleuThrAlaAlaHisCysLeuGlyProAspVal 80
Db 188 GGCCTCCCTATCCACCCAGGGGTGCTGACCGGACGACATGCTGGACCCGAGGATC 247
Qy 81 LysAspLeuAlaThrIleuArgValGlnIleuArgGlnGlnHisIleuTyrTrpGlnAspGln 100
Db 248 AAGGATTCGGCCCGCTCAGGGTGCACCTGGGGAGCGACCTTACTACAGGACCAAG 307
Qy 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyVal 120
Db 308 CTGCTGCGCGGTGAGGAGGATCTGTGACCCACAGTTCACCCGCCACAGTTCGAGCG 367
Qy 121 AspIleAlaLeuLeuGlnIleuGlnIleuProValAsnIleSerSerArgValHisThrVal 140
Db 368 GACATTCGCCCTCTGAGCTGGAGGAGCGCGGTGAAGCTTCACAGCCACGTCACCGGTC 427
Qy 141 MetLeuProProAlaSerGlnuThrPheProGlyMetProCysTrpValThrGlyTrp 160
Db 428 ACCCTGCCCTGCTCAGAGACCTTCCCGGGGATGCCGTGCTGGATCAGCTGGCTGG 487
Qy 161 GlyAspValAspAsnAspGluProLeuProProPheProLeuIleGlnValIlyVal 180
Db 488 GCGGATGGGACAAAGATGAGCGCCTCCACCGCCATTCCTGAAACAGGTGAAGGTC 547
Qy 181 ProIleMetGluAsnHisIleCysAspAlaIlysrYrHisIleuGlyAlaIlyrThrGlyAsp 200
Db 548 CCCATTAATGAAAAACAATTTGTGACCGAANAATTCACCTTGGCGCTTACAGGGAGAC 607
Qy 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
Db 608 GACGTCGGGATGCTCGAGACACATGCTGTGTGCGGGGAAACCCCGGAGGACTCATGTC 667
Qy 221 LysGlyAspSerGlyGlyProLeuValCysIlyValIleGlyThrTrpLeuGlnAlaGly 240
Db 668 CAGGCGGACTCCGGAGCGCCCTGTGTGTGCAAGGATGGACCTGCTGACAGCGGGG 727
Qy 241 ValValSerTrpAspGluGlyCysAlaGlnProAsnAspProGlyIleIlyrThrArgVal 260
Db 728 GTGTCAGCTGGGGGAGGGCTGTGCCCCAGCCCAACCGGCTTACACCCGCTGTC 787
Qy 261 ThrTyrTyrLeuAspTrpIleHisIlyrValProIlyIlyrPro 275
Db 788 ACCTACTACTTGTGACTGGATCCACCATATGTCACCAAAAAGCCG 832

RESULT 10
US-10-956-157-2444
; Sequence 2444, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatencIn version 3.2
; SEQ ID NO 2444
; LENGTH: 1148
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2444

Alignment Scores:
Pred. No.: 3.81e-158 Length: 1148
Score: 1405.00 Matches: 253
Percent Similarity: 94.91% Conservative: 8

```


Db 1031 ACCCTGCCCCCTGCTCAGACGACCTTCCCCGGGGATGCGCTGGTCACTGGCTGG 1090
 QY 161 G1YASpValAspAsnAspGluProLeuProProPheProLeuYsgInValIysVal 180
 Db 1091 GGGGATGTGGACATATGATGAGCGCCCTCCACCGCCATTTCTCTGAAAGATGMAAGGTC 1150
 QY 181 ProIleMeGluAsnHisIleCysAspAlaIysTyrHisIleGluIValIaTyrThrGlyAsp 200
 Db 1151 CCATATATGGAAACCAACATTTGTGACGGMAAATACACCTTGGCCCTTACACCGGAGAGC 1210
 QY 201 AppValArgIleIleLeuArgAspAspMetLeuCysAlaG1YAsnSerGlnArgAspSerCys 220
 Db 1211 GACGTCCGCAATCGTCCGTGACAGATGCTGTGTGCGGGGAAACACCGGGAGGACTATGTC 1270
 QY 221 LysG1YAspSerG1Yg1YProLeuValCysIysValAsnG1YThrTrpLeuGlnAlaG1Y 240
 Db 1271 CAGGGGAGACTCCGGAGAGGGCCCTGGTGTGCAAGGTGATGGCACTGGGTGGAGGGGGC 1330
 QY 241 ValValSerTrpAspGluG1YCysAlaGlnProAsnArgProG1YIleTyrThrArgVal 260
 Db 1331 GTGGTCAGATGGGGAGGGGCTGTGCCCCAGCCCAACCGGCTGGCACTTACACCCCGTGC 1390
 QY 261 ThrTyrTrpLeuAspTrpIleHisIleTyrValProIysIysPro 275
 Db 1391 ACCTACTACTTGGACTGATCCACCACTATGTCCCAAAAAGCCG 1435
 RESULT 12
 US-09-954-456-2126 / Sequence 2126, Application US/09954456
 / Patent No. US20020115057A1
 / GENERAL INFORMATION:
 / APPLICANT: Young, Paul
 / TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 / FILE REFERENCE: 689290-76
 / CURRENT APPLICATION NUMBER: US/09/954,456
 / PRIOR FILING DATE: 2001-09-18
 / PRIOR APPLICATION NUMBER: US/60/233,617
 / PRIOR FILING DATE: 2000-09-18
 / PRIOR APPLICATION NUMBER: US/60/234,052
 / PRIOR FILING DATE: 2000-09-20
 / PRIOR APPLICATION NUMBER: US/60/234,923
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,134
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,637
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,638
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,711
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,720
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,840
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,863
 / NUMBER OF SEQ ID NOS: 2276
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 2126
 / LENGTH: 1081
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-09-954-456-2126

US-09-598-982c-52 (1-275) x US-09-954-456-2126 (1-1081)
 QY 9 LeuProValIleValAsnSerArgAlaTyrAlaIleProValIleValIleGlnGln 28
 Db 2 CTGGCCGCTCGGAGGCGGGCCCTTACGGCGGCCCTGCCCCAGGCGCCCTGAGCGCA 61
 QY 29 AlaGlyIleValGlyGlnGlnAlaProArgSerIysTrpProTrpGlnIValSerIeu 48
 Db 62 GTGGGCAATCGTGGGGGTGAGAGGCCCCCAAGCAAGTGGCCCTGGCAGGTGAGCTGTG 121
 QY 49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
 Db 122 AGAGTCCCGGACCAATCTGATGATGACTTTCGGGGGGCTCCCTATCCACCCCAAGTGG 181
 QY 69 ValLeuThrAlaAlaHisCysIleuGlyProAspValIysAspLeuAlaThrLeuArgVal 88
 Db 182 GTGCTGACCCGAGCGCACTGGCGTGGGACCCGAGCGTCAAGATGTGGCCGCTCAGGGGTG 241
 QY 89 GlnLeuArgGluGlnHisIleuTyrTyrGlnIleAspGlnLeuLeuProValSerArgIleIle 108
 Db 242 CAACCTGCGGAGCAGCACTTACCAAGGACCGACTGCTGCGCGGTGAGCAGGATTC 301
 QY 109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuIleuGluIle 128
 Db 302 GTGCACCCAGACTTCTTACACCGCCAGATCGGAGCGGACATCGCCCTGTGAGGCTGGAG 361
 QY 129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
 Db 352 GAGCGCGTGGAGGATGCTCCAGCCAGCTCCACCGGTCACTGCGCCCTCAGAGAGCC 421
 QY 149 PheProProGlyMetProCysETrpValIleThrGlyIysAspValAspAsnAspGluPro 168
 Db 422 TTCGCCCGGGGATCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
 QY 169 LeuProProPheProLeuIleuGlnValIysValProIleMetGluAsnHisIleCys 188
 Db 482 CTCGCCACCGCCATTTCTCTGTAAGAGGTGAAGGTCCCAATATGAAACCAACATTTGT 541
 QY 189 AspAlaIysTyrHisIleuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
 Db 542 GACGGAAATATCACCTTGGCGCTTACCGGGAAGCAGCGTCCGATGCTCCGTGACGAC 601
 QY 209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysIysGlyAspSerGlyIProIeu 228
 Db 602 ATGCTGTGTGCGGGAAACCCCGGAGACTCATGCGGGCGACTCCGGAGGGCCCTGTG 661
 QY 229 ValCysIysValAsnG1YThrTrpLeuGlnAlaG1YValIleSerTrpAspG1UG1YCys 248
 Db 662 GTGTGCAAGGTGATGGAACCTGCTGCAAGCGGGCGGTGCTCAGCTGGGGGAGGGGCTGT 721
 QY 249 AlaGlnProAsnArgProG1YIleTyrThrArgValIleTyrTyrIleuAspTrpIleHis 268
 Db 722 GCCGAGCCCAACCGGCTGGGATTCAGACCCGTTGCACTACTTGGACTGATCCAC 781
 QY 269 HisTyrValProIysIysPro 275
 Db 782 CACTATGTCCCAAAAAGCCG 802
 RESULT 13
 US-09-960-706-680 / Sequence 680, Application US/09960706
 / Publication No. US20030134280A1
 / GENERAL INFORMATION:
 / APPLICANT: Mungert, William E.
 / TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasi
 / FILE REFERENCE: 44921-5029-01US
 / CURRENT APPLICATION NUMBER: US/09/960,706
 / PRIOR FILING DATE: 2001-09-24
 / PRIOR APPLICATION NUMBER: 60/223,323
 / PRIOR FILING DATE: 2000-08-07
 / PRIOR APPLICATION NUMBER: 09/873,319
 / PRIOR FILING DATE: 2001-06-05

Alignment Scores:
 Pred. No.: 7,36e-157 Length: 1081
 Score: 1394.00 Matches: 249
 Percent Similarity: 95.88% Conservative: 7
 Percent Local Similarity: 93.26% Mismatches: 11
 Query Match: 92.20% Indels: 0
 DB: 9 Gaps: 0

NUMBER OF SEQ ID NOS: 1124
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 680
 LENGTH: 1081
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US200301343280A1 M33493
 US-09-960-706-680

Alignment Scores:
 Pred. No.: 7,36e-157 Length: 1081
 Score: 1394.00 Matches: 249
 Percent Similarity: 95.88% Conservative: 7
 Best Local Similarity: 93.26% Mismatches: 11
 Query Match: 92.20% Indels: 0
 Gaps: 0

US-09-598-982C-52 (1-275) x US-09-960-706-680 (1-1081)
 QY 9 LeuPProValLeuAlaSerArgAlaTyrAlaAlaProValGlnAlaLeuGlnGln 28
 DB 2 GTGCCCGCTCCGCGGAGCGCCCTTACCGCGCCCTCCAGCCAGCCCTCGAGCGGA 61
 QY 29 AlaGlyIleValGlyGlnGlnAlaProArgSerIleTyrProTyrGlnValSerIleu 48
 DB 62 GTGGGCAATCGTTGGGGGTCAGAGGCCCCCGAGAGAGAGTGGCCCTGGAGGACCTG 121
 QY 49 ArgValAlaArgAspArgTyrTrpMetHisPheCysGlyGlySerIleuIleHisProGlnTyr 68
 DB 122 AGAGCCCGGAGCCGATACGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
 QY 69 ValLeuThrAlaAlaHisCysLeuGlyProAspValIleAspLeuAlaThrLeuArgVal 88
 DB 182 GTGCTGACCGCGCCGACCTGCTGGAGCCGTCAGAGATGATGATGATGATGATGATG 241
 QY 89 GlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGlnIleuLeuProValSerArgIleIle 108
 DB 242 CAACCTGCGGAGACGACCTTCTACGACGACGACGACGACGACGACGACGACGACGACG 301
 QY 109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGlnIleuGln 128
 DB 302 GTGGCCCGCAGAGTCTTACACCGCCAGATCGGAGGAGATCGCCCTGCTGGAGCTGG 361
 QY 129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
 DB 362 GAGCCCGGTGAAGGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 421
 QY 149 PheProProGlyMetProCysTyrValIleIleGlnThrGlyAspValAspAspAspGluPro 168
 DB 422 TTCCCCCGGGGATGCGGTGGTCACTGGAGCTGGAGGATGATGATGATGATGATGATG 481
 QY 169 LeuProProProPheProLeuIleValIleValProIleMetGluAsnHisIleCys 188
 DB 482 CTCCACCGCCATCTTCCTGAAAGCAAGGATGATGATGATGATGATGATGATGATGATG 541
 QY 189 AspAlaIleTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
 DB 542 GACGGAAATATACACCTTGGCCCTTACACCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 QY 209 MetLeuCysValAlaGlnAsnSerGlnArgAspSerCysValGlyAspSerGlyGlyProLeu 228
 DB 602 ATGCTGTGTGGGAGACCCCGAGAGGACTATAGCCAGAGGAGACTCCGAGAGGAGCCCTG 661
 QY 229 ValCysLeuValAsnGlyThrTrpLeuGlnAlaGlyValIleSerTrpAspIleuGlyCys 248
 DB 662 GTGTGCAAGGTGAATGGGCACTGGCTGCAAGGCGGGGGGTGTGAGTGGGGGCGAGGCTGT 721
 QY 249 AlaGlnProAspArgProGlyIleTyrThrArgValIleTyrTyrLeuAspTrpIleHis 268
 DB 722 GCCCAGCCCAACCGCGCTGGACATCTACACCCCGTGTCACTTACTGAGATCGATCCAC 781
 QY 269 HisTyrValProLeuIleValPro 275

DB 782 CACTATGTCGCCCAAAAAGCCG 802

RESULT 14
 US-09-873-319-427 Application US/09873319A
 ; Sequence 427, Publication No. US20030134324A1
 ; Publication No. US20030134324A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mungert, William E.
 ; APPLICANT: Kulkarni, Prakash
 ; APPLICANT: Getzenberg, Robert H.
 ; APPLICANT: Waga, Iwao
 ; APPLICANT: Yamamoto, Jun
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
 ; TITLE OF INVENTION: Hypertlasia Using Gene Expression Profiles
 ; FILE REFERENCE: 44921-5029-US
 ; CURRENT FILING DATE: 2001-06-05
 ; EARLIER FILING DATE: 2000-08-07
 ; EARLIER APPLICATION NUMBER: US 60/223,323
 ; NUMBER OF SEQ ID NOS: 755
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 427
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
 US-09-873-319-427

Alignment Scores:
 Pred. No.: 7,36e-157 Length: 1081
 Score: 1394.00 Matches: 249
 Percent Similarity: 95.88% Conservative: 7
 Best Local Similarity: 93.26% Mismatches: 11
 Query Match: 92.20% Indels: 0
 Gaps: 0

US-09-598-982C-52 (1-275) x US-09-873-319-427 (1-1081)
 QY 9 LeuPProValLeuAlaSerArgAlaTyrAlaAlaProValGlnAlaLeuGlnGln 28
 DB 2 GTGCCCGCTCCGCGGAGCGCCCTTACCGCGCCCTCCAGCCAGCCCTCGAGCGGA 61
 QY 29 AlaGlyIleValGlyGlnGlnAlaProArgSerIleTyrProTyrGlnValSerIleu 48
 DB 62 GTGGGCAATCGTTGGGGGTCAGAGGCCCCCGAGAGAGAGTGGCCCTGGAGGACCTG 121
 QY 49 ArgValAlaArgAspArgTyrTrpMetHisPheCysGlyGlySerIleuIleHisProGlnTyr 68
 DB 122 AGAGCCCGGAGCCGATACGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
 QY 69 ValLeuThrAlaAlaHisCysLeuGlyProAspValIleAspLeuAlaThrLeuArgVal 88
 DB 182 GTGCTGACCGCGCCGACCTGCTGGAGCCGTCAGAGATGATGATGATGATGATGATG 241
 QY 89 GlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGlnIleuLeuProValSerArgIleIle 108
 DB 242 CAACCTGCGGAGACGACCTTCTACGACGACGACGACGACGACGACGACGACGACGACG 301
 QY 109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGlnIleuGln 128
 DB 302 GTGGCCCGCAGAGTCTTACACCGCCAGATCGGAGGAGATCGCCCTGCTGGAGCTGG 361
 QY 129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
 DB 362 GAGCCCGGTGAAGGTCTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 421
 QY 149 PheProProGlyMetProCysTyrValIleIleGlnThrGlyAspValAspAspAspGluPro 168
 DB 422 TTCCCCCGGGGATGCGGTGGTCACTGGAGCTGGAGGATGATGATGATGATGATGATG 481
 QY 169 LeuProProProPheProLeuIleValIleValProIleMetGluAsnHisIleCys 188

```

DB 482 CTCGCCACCCCAATTCCTCTGAAAGCAGTGAAGGTCCTCCCATATGGAACCAACCAATTTGT 541
QY 189 AAPPALALVRYRHISLEUGLYVALATYRTRHGIYASPAAPVALARGILEILHARGASPAAP 208
DB 542 GACGGAAATATACCACTTGTGGCCCTTACCGGAGACGACGCTCCGATTCGTCCGTGACGAC 601
QY 209 MetLeuCYsAlaGlyAsnSerGlnArgAspSerCYsAluGlyAspSerGlyProLeu 228
DB 602 ATGCTGTGTGGCCGGGAAACCCCGAGAGGACTCATGTCAGAGCCGACTCCGAGGGCCCTG 661
QY 229 ValCYsIysValAsnGlyThrTrrPLeuGlnAlaGlyValValSerTrrAspGluGlyCys 248
DB 662 GTGTGCAGAGTGAATGTGACCTGTGCGAGCGGGGCTGCTCACTGGGGCGAGGCTGT 721
QY 249 AlaGlnProAsnArgProGlyIleYrThrArgValThrTYrTrrLeuAspTrrIleHis 268
DB 722 GCCCAGCCCAACCGGCTTGGCATCTACACCCGTGTGACCTTACTACTTGGACGTGATCCAC 781
QY 269 HisTYrValProIyAspPro 275
DB 782 CACTATGTCCCAAAAAAGCCG 802

```

```

RESULT 15
US-09-873-367C-155
/ Sequence 155, Application US/09873367C
/ Publication No. US20030165839A1
/ GENERAL INFORMATION:
/ APPLICANT: Young, Paul
/ APPLICANT: Soggett, Daniel
/ APPLICANT: Endress, Gregory
/ APPLICANT: Augustus, Meena
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Carter, Kenneth
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
/ FILE REFERENCE: 689290-64
/ CURRENT APPLICATION NUMBER: US/09/873,367C
/ PRIOR FILING DATE: 2003-04-29
/ PRIOR APPLICATION NUMBER: U.S. 60/236,891
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: U.S. 60/236,842
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: U.S. 60/244,867
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: U.S. 60/245,084
/ PRIOR FILING DATE: 2000-11-01
/ NUMBER OF SEQ ID NOS: 1067
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 155
/ TYPE: DNA
/ LENGTH: 1081
/ ORGANISM: Homo sapiens
US-09-873-367C-155

```

```

Alignment Scores:
Pred. No.: 7.36e-157 Length: 1081
Score: 1394.00 Matches: 249
Percent Similarity: 95.884 Conservative: 7
Best Local Similarity: 93.268 Mismatches: 11
Query Match: 92.208 Indels: 0
DB: 10 Gaps: 0

```

```

US-09-598-982C-52 (1-275) x US-09-873-367C-155 (1-1081)
QY 9 LeuProValLeuAlaSerArgAlaTYrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
DB 2 CTGCCCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 61
QY 29 AlaGlyIleValGlyIysGlnGlnAlaProArgSerIysTrrProTrrGlnValSerLeu 48
DB 62 GTGGGATTCGTTGGGGGTCAAGAGGCCCCCAAGAGCAAGTGGCCCTTGGACGATGACCTG 121

```

```

QY 49 ArgValAlaArgAspArgTYrTrrPMeChIspheCYsGlyGlySerLeuIleHisProGlnTrr 68
DB 122 AGAGTCCCGCAGCCGATACTGTGATGCACTTGTGGGGGGGCTCCCTCATCCACCCCGAGGG 181
QY 69 ValLeuThrAlaAlaHisCYsLeuGlyProAspValIysAspLeuAlaThrLeuArgVal 88
DB 182 GTGTGACCGGACCGCACTGTGGGACCGGACGTCAAGATCTGSCCGCCCTCAGGGTGG 241
QY 89 GlnLeuArgGlnGlnHisLeuTYrTYrGlnAspGlnLeuLeuProValSerArgIleIle 108
DB 242 CAACGTGGGGAGCAGACACTCTACTACAGAGGACGAGCTGTGGCGGCTGAGGATCAATC 301
QY 109 ValHisProGlnPheTrrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGlnLeuGlu 128
DB 302 GTGCACCCCAAGTTCCTTACACCCCGCATCGAGGAGGACATCGCCCTGTGAGCTGGAG 361
QY 129 GlnProValAsnIleSerSerArgValHisPheValIleLeuLeuProAlaSerGluThr 148
DB 362 GAGCCGTTGAAAGTTCACAGCCAGGTCACAGGTCACCCCTGCTGCTGACAGAGACC 421
QY 149 PheProProGlyMetProCYsTrrValThrGlyTrrGlyAspValAspAsnAspGluPro 168
DB 422 TTCGCCCGGGGATGCGGTGGGTCACTGGCTGGGGGATGTCAGATGATGATGAGCGC 481
QY 169 LeuProProPheProLeuLeuValIysValProIleMetGlnAsnHisIleCys 188
DB 482 CTCGCCACCGCCATTTCTCTGTGAGGAGGATGAAAGTCCCAATATGGAACCAACCAATTTGT 541
QY 189 AAPPALALVRYRHISLEUGLYVALATYRTRHGIYASPAAPVALARGILEILHARGASPAAP 208
DB 542 GACGGAAATATACCACTTGTGGCCCTTACCGGAGACGACGCTCCGATTCGTCCGTGACGAC 601
QY 209 MetLeuCYsAlaGlyAsnSerGlnArgAspSerCYsAluGlyAspSerGlyProLeu 228
DB 602 ATGCTGTGTGGCCGGGAAACCCCGAGAGGACTCATGTCAGAGCCGACTCCGAGGGCCCTG 661
QY 229 ValCYsIysValAsnGlyThrTrrPLeuGlnAlaGlyValValSerTrrAspGluGlyCys 248
DB 662 GTGTGCAGAGTGAATGTGACCTGTGCGAGCGGGGCTGCTCACTGGGGCGAGGCTGT 721
QY 249 AlaGlnProAsnArgProGlyIleYrThrArgValThrTYrTrrLeuAspTrrIleHis 268
DB 722 GCCCAGCCCAACCGGCTTGGCATCTACACCCGTGTGACCTTACTACTTGGACGTGATCCAC 781
QY 269 HisTYrValProIyAspPro 275
DB 782 CACTATGTCCCAAAAAAGCCG 802

```

```

Search completed: August 28, 2005, 01:30:05
Job time : 622 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 23:30:16 ; Search time 186 Seconds
(without alignments)
2419.226 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1 MLSLLLLALPVLASRAVYAP.....IYTRVYVYLDWVHHVPRKP 275
Sequence: 1512

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODL=frame+_p2n_model -DEV=xlp
-Q=/cg2_1/USPRO_spool_P/US09598982/runat_25082005_165826_1568/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastcap -SUPFIX=xrml -MINMATCH=0.1 -LOOPCL=0
-LDOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982@cgn_1_1_105@runat_25082005_165826_1568 -NCPD=6 -ICPD=3
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfillseq1.seq:*

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains sequence alignment data for various database entries.

Table with columns: Hit ID, Score, Query Match, Length, DB ID, Description. Contains sequence alignment data for various database entries.

ALIGNMENTS

RESULT 1
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wolf, Greenfield & Sacks, P. C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016_366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/037_090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441

TELEX:
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1154 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-016-366A-16

Alignment Scores:
 Pred. No.: 1,56e-156 Length: 1154
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 2 0

US-09-598-982C-52 (1-275) x US-09-016-366A-16 (1-1154)

1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 18 ATGCTGAGGCTGCTGCTGGGGCTGCCCTGCTGGGAGCCGGCCCTAAGCGGGCCCT 77
 21 AlaProValGlnAlaLeuGlnGlnAlaGlyTlleValGlyGlnGlnAlaProArgSer 40
 78 GCCCCAGTCCAGGCTGCGCAAGCGGGGTATGTCGGGGGTCAAGGAGCCCCCAAGAAC 137
 41 LysTrpProTrrGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 138 AAGTGGCCCTGGAGGGTGAAGCTGAGAGTCCGGAGCCGAACTGGAAAGCACTTCTGGGG 197
 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 198 GGTTCCTCATCAACCCCAAGTGGGGTGTACCGGGGGCGACTGGGGACCGGACGTC 257
 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGln 100
 258 AAGGATCTGGCCACCTCCAGGGGTGCAACTGCGGAGACAGCACTCTACTAACCAGACCA 317
 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnHrgVala 120
 318 CTGCTGCGAATCGAGGATCGAGTCAATCGACCCCAAGTTCACATCCAGACTGGAGCG 377
 121 AspIleAlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrVal 140
 378 GATATGCGCCCTGCTGAGCTGGAGGAGCCCGTGAACATCTCCAGCCCGCTCCACAGCG 437
 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrr 160
 438 ATGCTGCCCCCTCCCTCGGAGACCTTCCCCCGGGGATGCGCTGGGTCACTGGCTGC 497
 161 GlyAspValAspAsnAspGluProLeuProProPheProLeuValysGlnValLysVal 180
 498 GGGAGTGTGACATGATGATGAGCCCTCCACCGCATTTCCCTGAAAGCAAGGAAAGTTC 557
 181 ProIleMetGlnAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrTrpGlyAsp 200
 558 CCAATATGAAAACCAATTTGTGAGCGCAAAATACCACTTGGCCGCTAACAAGGAGAC 617
 201 AspValArgIleIleArgAspAspMetLeuValGlyAsnSerGlnArgAspSerCys 220
 618 GAGGTCGCGCATCTCCGTGACGACATGCTGTGCGGGGAAACAGCCAGAGGGGCTCTGC 677
 221 LysGlyAspSerIleGlyProLeuValCysLysValAsnGlyTyrTrpLeuGlnAlaGly 240
 678 AAGGGCCACTCTGAGAGGGCCCTGGTGTGCAAGGTGATGACCTGGCTAAGGGGGGG 737
 241 ValValSerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrTrpArgVal 260
 738 GTGGTACGCTGGAGCAAGGGCTGTGGCCCAACCGGCTCTGACATCCCGCTGTTC 797
 261 ThrTyrTyrLeuAspTrpIleHisIleTyrValProLysLysAspPro 275

798 ACCTACTGACTGGACTGCATCCACCACTATGTCCCAAAAAGCCG 842

RESULT 2
 US-08-978-404B-11
 Sequence 11, Application US/08978404B
 Patent No. 5968782
 GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 TITLE OF INVENTION: FIBRINOGEN
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441

TELEX:
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1154 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-978-404B-11

Alignment Scores:
 Pred. No.: 1,56e-156 Length: 1154
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 2 0

US-09-598-982C-52 (1-275) x US-08-978-404B-11 (1-1154)

1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 18 ATGCTGAGGCTGCTGCTGGGGCTGCCCTGCTGGGAGCCGGCCCTAAGCGGGCCCT 77
 21 AlaProValGlnAlaLeuGlnGlnAlaGlyTlleValGlyGlnGlnAlaProArgSer 40
 78 GCCCCAGTCCAGGCTGCGCAAGCGGGGTATGTCGGGGGTCAAGGAGCCCCCAAGAAC 137
 41 LysTrpProTrrGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 138 AAGTGGCCCTGGAGGGTGAAGCTGAGAGTCCGGAGCCGAACTGGAAAGCACTTCTGGGG 197
 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 198 GGTTCCTCATCAACCCCAAGTGGGGTGTACCGGGGGCGACTGGGGACCGGACGTC 257

QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGln 100
 DB 258 AAGATCTGGCCACCTTCAAGGTGCACTGGCGGAGGAGCACTTACTACACAGGACGAG 317
 QY 101 LeuLeuProValSerArgLleIleValHisProGlnPheTyrLleIleGlnThrGlyAla 120
 DB 318 CTGGCTGGCAGGTGAGGAGGATCATCTGTCACCCACAGTTTCAATCATCATCATGAGGAGCG 377
 QY 121 AspIleAlaLeuLeuGlnLeuGlnGlnIleuProValAsnIleSerSerArgValHisThrVal 140
 DB 378 GATATTCGGCCCTGCTGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGGCTCCACCGGTC 437
 QY 141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTyrValThrGlyTyrP 160
 DB 438 ATGCTGCCCCCTGCTCGGAGACCTTCCCCCGGGGAGTCCGCTGGGTCACTGGCTGG 497
 QY 161 GlyAspValAspAsnAspGlnuProLeuProProProPheProLeuLysGlnValIlyVal 180
 DB 498 GCGGATGTGGCAAAATGATGAGCCCTCCACCGCCATTTCCCTGAAAGGAGGTGAAGGTC 557
 QY 181 ProIleMetGluAsnHisIleCysAspAlaIlyTyrHisIleuGlyValATyrThrGlyAsp 200
 DB 558 CCATAAATGGAAAACCACTTTGTGACCGAAAATACCACTTGGCGCTTACACCGGAGAGC 617
 QY 201 AspValArgLleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 618 GACGTCCCGCATCATCCGTGAGCAGCATGCTGTGTGCGGGGAGACGACAGGAGGACTCTTCC 677
 QY 221 LysGlyAspSerGlyGlyProLeuValCysIlyValAsnGlyTyrThrTyrLeuGlnAlaGly 240
 DB 678 AAGGCGCACTGTGAGGGGCCCCCTGGTGTGCAAGTGAATGTGCACTGGGCTACAGCGGGG 737
 QY 241 ValIleSerTyrAspGlnGlyCysAlaGlnProAsnAspProGlyTyrIleTyrThrTyrVal 260
 DB 738 GTGGTCAGCTGGAGGAGGAGGCTGTGCCCGCCAAACCGGCTGGAGATCTACACCCGCTTC 797
 QY 261 ThrTyrTyrLeuAspTyrIleHisIlyTyrValProLysIlyPro 275
 DB 798 ACCTACTACTTGGACTGGATCCACCATATGTCTCCCAAAAAGCCG 842

RESULT 3
 US-09-016-366A-18
 ; Sequence 18, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Wolf, Greenfield & Sacke, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,366A
 ; FILING DATE: January 30, 1998
 ; CLASSIFICATION: 530
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/037,090
 ; FILING DATE: 05-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7093

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELERX:
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-016-366A-18
 Alignment Scores:
 Pident: 1.93e-144 Length: 1137
 Score: 1402.00 Matches: 252
 Percent Similarity: 95.24% Conservative: 8
 Best Local Similarity: 92.31% Mismatches: 13
 Query Match: 92.72% Indels: 0
 DB: Gaps: 0
 US-09-598-982C-52 (1-275) x US-09-016-366A-18 (1-1137)
 QY 3 SerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaPro 22
 DB 3 AATCGCTGCTGGCTGGGCTGCCCTCCGCTCGGCGAGCCGCGCTACGCGGCCCCCTGCCCCA 62
 QY 23 ValGlnAlaLeuGlnGlnAlaGlyTyrIleValGlyGlyGlnGlnAlaProArgSerIlyTyr 42
 DB 63 GCGCCAGGCGCTGCGAGCGGTGGCATGTCGGGGGTGACGAGGCGCCCGAGGAAAGTGG 122
 QY 43 ProTyrGlnValSerLeuArgValArgAspArgTyrTyrMetHisIleCysGlyGlySer 62
 DB 123 CCTTGGCAGGTGAGCTGAGATCCAGCGCCCATTCAGATGCACTTCTGGGGGGCTTCC 182
 QY 63 LeuIleHisProGlnTyrValIleuThrAlaAlaHisCysLeuGlyProAspValIlyAsp 82
 DB 183 CTCATCCACCCCGAGTGGGTCTGACCGCAGCGCATGCGTGGGAGCCGGAAGTCAAGGAT 242
 QY 83 LeuAlaThrLeuAlaGValGlnLeuArgGlnGlnHisIleuTyrTyrGlnAspGlnLeuLeu 102
 DB 243 CTGGCGCCCTCCAGAGGTGCAACCTCGGGAGGAGCACTCTACTACCGAGCACAGCTGTG 302
 QY 103 ProValSerArgLleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIle 122
 DB 303 CCGGTGAGGAGATATGTCACCCACAGTTTCTACACCGCCCAAGATGGAGCGGACATC 362
 QY 123 AlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrValMetLeu 142
 DB 363 GCCCTGCTGGAGCTGGAGGAGCCGGTGAAGCTTCCAGCCACGTCACAGCGGTACCCCTG 422
 QY 143 ProProAlaSerGlnuThrPheProProGlyMetProCysTyrTyrValIlyTyrGlyAsp 162
 DB 423 CCCCCCTGCTGAGGAGACTTCCCGGGGAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 482
 QY 163 ValAspAsnAspGlnuProLeuProProPheProLeuIlyGlnValIlyValProIle 182
 DB 483 GTGGCAATGATGAGCCGCTCCACCGCCATTTCTTGAAGCGAGGTGAAGTCCCATTA 542
 QY 183 MetGlnAsnHisIleCysAspAlaIlyTyrHisIleuGlyAlaTyrThrGlyAspAspVal 202
 DB 543 ATGGAANAACCAATTTGTGAGGCAAAATATACCACTTGGCGCTACAGGGAGAGAGAGTC 602
 QY 203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnAspSerCysIlyGly 222
 DB 603 CGCATTCGCTCGTGAACATCTGTGTGCGCGGGAACCCCGAGGAGCTCAATGCGCAAGGCG 662
 QY 223 AspSerGlyGlyProLeuValCysIlyValAsnGlyTyrThrTyrLeuGlnAlaGlyValIly 242
 DB 663 GACTCCGAGGAGGCGCCCTGGTGTGCAAGGTGAATGGACCTTGGCTGCAAGGGGGGCTGGTC 722
 QY 243 SerTyrAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValIlyTyr 262

DB 723 AGCTGGGGGAGGAGGCTGTGTGCCAGCCCAACCGGCGCTGGACCTACACCCGCTGTCACTTAC 782
 QY 263 TYLEUASPTTPILEHISHISTYRVALPROLYSLVSPRO 275
 DB 783 TACTTGGACTGGATCCACCACCTATGTCTCCCAAAAAGCCG 821

RESULT 4
 US-08-978-404B-13
 ; Sequence 13, Application US/08978404B
 ; Patent No. 5968782

GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEINASE THAT CLEAVES
 TITLE OF INVENTION: FIBRINOGEN
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978,404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELETYPE:
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Alignment Scores:
 Pred. No.: 1,93e-144 Length: 1137
 Score: 1402.00 Matches: 252
 Percent Similarity: 95.24% Conservative: 8
 Best Local Similarity: 92.31% Mismatches: 13
 Query Match: 92.72% Indels: 0
 DB: 2 Gaps: 0

US-09-598-982c-52 (1-275) x US-08-978-404B-13 (1-1137)
 QY 3 SerLeuLeuLeuLeuLalLeuProValLeuLalSerArgLalArgLalAlaLalProAlaPro 22
 DB 3 AATCTGTCTGTCTGGGCTGCGCCGCTCTGGCGAGCCGCCCTTACCCCGGCGCTGCGCC 62
 QY 23 ValGlnAlaLeuGlnGlnGlnAlaGlyValLeuGlyGlnGlnGlnAlaProArgSerLysTrp 42
 DB 63 GGCAGGCGCCCTGGAGGAGTGGGCGATGTGTGGGGGTGAGGAGCCGCCAGAGCAAGTGG 122
 QY 43 ProTrpGlnValSerLeuArgValArgAspArgTrpTrpMetHisPheCysGlyGlySer 62
 DB 123 CCTGGAGGAGTGGAGCTGAGAGTCCAGCGCCCATTACTGGATGACATTTCTGGGGGCTCC 182
 QY 63 LeuIleHisPheProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82

DB 183 CTATCCACCCCACTGGAGTGTGACCCGACCGCACTGCGGGAGCCGAGCTCAAGGAT 242
 QY 83 LeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuLeuTrpGlnAlaAspGlnLeuLeu 102
 DB 243 CTGGCCCGCCCTCAGGGGTGCAACTGGGGAGCAGCAGCCTTACTACAGGAGCAGGCTGTG 302
 QY 103 ProValSerArgIleIleValHisPheGlnPheTrpIleIleGlnThrGlyAlaAspIle 122
 DB 303 CCGGTCAAGAGGATGATGTCGACCCACAGTTCTTACCCGCCAGATGGAGCGGAGCAATC 362

QY 123 AlaLeuLeuGlnLeuGlnGluProValAsnIleSerSerArgValHisThrValMetLeu 142
 DB 363 GCCCTGCTGGAGCTGAGAGCGGGTGAAGCTTCCAGCCACGTCACACAGGTCACCTTG 422
 QY 143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
 DB 423 CCCCCTGCTCAGAGACTTCCCGCCCGGGAGTCCGCTGCTGGTCACTGGCGGGCGAT 482

QY 163 ValAspAsnAspGluProLeuProProPheProLeuLysGlnValLysValProIle 182
 DB 483 GTGACAAATGATGAGCGCTCCACCCGCAATTTCTTGAAGCAGGTGAAGGTCCCAATA 542
 QY 183 MetGluAsnHisIleCysAspAlaIleTyrHisIleuGlyAlaTyrThrGlyAspAspVal 202
 DB 543 ATGGAAACCAACATTTGTGACGCAAAATVACACTTGGCGCTTACACCGGAGAGCAGCTC 602

QY 203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLeuGly 222
 DB 603 CGCATCGTCCGTCAGAGCATGTGTGTCCCGGAAACACCCGAGGAGATCATGTCAGAGGG 662
 QY 223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
 DB 663 GACTCCGGAGGGCCCTGAGTGTGCAAGGTGAATGGCACCTGGCTGCAAGCGGGCGTGGTTC 722

QY 243 SerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
 DB 723 ACCTGGGGCGAGGGCTGTGCCAGCCCAACCGGCGCTGGCACTTACACCCGCTGTCACTAC 782
 QY 263 TYLEUASPTTPILEHISHISTYRVALPROLYSLVSPRO 275
 DB 783 TACTTGGACTGGATCCACCACCTATGTCTCCCAAAAAGCCG 821

RESULT 5
 US-09-016-366A-20
 ; Sequence 20, Application US/09016366A
 ; Patent No. 5955431
 GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEINASE PEPTIDE
 TITLE OF INVENTION: INHIBITORS
 NUMBER OF SEQUENCES: 65
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,366A
 FILING DATE: January 30, 1998
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,090
 FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:

Db 122 TGGCCCTGGCAGGTGAGCTGAGAGTCCAGCGCCCACTACTGATGCACTTCTGGCGGGGCG 181
 QY SerLeu11ehisProGlnTrpValLeuThrAlaAlahisCysLeuGlyProAspVallys 81
 Db 182 TCCCTCATCCACCACCAGTGGGCTGACCGCAGCCCACTGGCTGGGAGCCGGAACG 241
 QY 82 APRLAVALATHRLEUARGVALGINDLEUARGGINDHISLEUTRYTRGINDASPGINDLEU 101
 Db 242 GATCTGGCCCGCCCTCAGGGGTGCAACTGCGGGAGCACAACCTCTACTACCAAGGACCCAGCTG 301
 QY 102 LeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAsp 121
 Db 302 CTGCGCGGTGAGCGAGATCATGTGACCCACAGTCTTACACCGGCCACATCGGAGCGGAGC 361
 QY 122 IleAlaLeuLeuGlnLeuGlnGluProValAsnIleSerSerArgValHisThrValMet 141
 Db 362 ATCGCCCTGTGGAGGTGAGGAGCGCGGTGAAAGTCTCCAGCCACAGCTCCACACCGGTCAAC 421
 QY 142 LeuProProAlaSerGluThrPheProProGlyMetProCysTrpValIleThrGlyTrpGly 161
 Db 422 CTGCCCCCTGCTCAGAGACCTTCCCCCGGGGATGCCGTGGCTGCTCACTGGCTGGGGCC 481
 QY 162 APVALASPAASNPGLUPROLEUPROPROPROPHLEUPLYGSLNVALYSPALPRO 181
 Db 482 GATGTGAAATGATGATGAGCGCCCTCCACCGCCATTTCTCTGAAAGAGGTGAAAGGTCCCC 541
 QY 182 IleMetGlnAsnHisIleCysAspAlaIleTyrTrhisIleuGlyAlaIleTyrTrhisIleu 201
 Db 542 ATAAATGAAACCAACATTTGTGTGACGCAAAATACACTTTGGGGCCCTTACACAGGGAGAGCAGC 601
 QY 202 ValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysIlys 221
 Db 602 GTCCCGCATGTCCGTGACGACATGTGTGTGCGCGGAGAACACCCGGAGGAGCATCATGCCAG 661
 QY 222 GlyAspSerGlyGlyProLeuValCysIlysValAsnGlyThrTrpLeuGlnAlaIleGlyVal 241
 Db 662 GCGGACTCCGAGAGGGCCCTGGTGTGACAGGTGAAAGTGAAGGCACTGGGCTGCAAGCGGCGGTG 721
 QY 242 ValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrTrhisIleuValThr 261
 Db 722 GTGAGTGGGGGGAGGGGCTGTGCCCAAGCCCAACCGGCTGGGACTCTACACCCGCTGTCAACC 781
 QY 262 TTTTTLTLEUASPTTPILEHISITRYVALPROLYSLEUPRO 275
 Db 782 TACTACTGTGACTGATCCACCACTATGTCCCCCAAAAAGCCG 823
 RESULT 7
 US-09-016-366A-22
 / Sequence 22, Application US/09016366A
 / Patent No. 5955431
 / GENERAL INFORMATION:
 / APPLICANT: Huang, Richard L.
 / TITLE OF INVENTION: MAST CELL PROTEASB PEPTIDE
 / TITLE OR INVENTION: INHIBITORS
 / NUMBER OF SEQUENCES: 65
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 / STREET: 600 Atlantic Avenue
 / CITY: Boston
 / STATE: MA
 / COUNTRY: U.S.A.
 / ZIP: 02210-2211
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Diskette
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: DOS
 / SOFTWARE: FASTSEQ For Windows Version 2.0
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/016,366A
 / FILING DATE: January 30, 1998
 / CLASSIFICATION: 530
 / PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 60/037,090
 / FILING DATE: 05-FEB-1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Plumer, Elizabeth R.
 / REGISTRATION NUMBER: 36,637
 / REFERENCE/DOCKET NUMBER: B0801/7093
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 617-720-3500
 / TELEFAX: 617-720-2441
 / TELEEX:
 / INFORMATION FOR SEQ ID NO: 22:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1081 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA
 / US-09-016-366A-22
 Alignment Scores:
 Pred. No.: 1,35e-143 Length: 1081
 Score: 1394.00 Matches: 249
 Percent Similarity: 95.88% Conservative: 7
 Best Local Similarity: 92.25% Mismatches: 11
 Query Match: 92.20% Indels: 0
 DB: 2 Gaps: 0
 US-09-598-982C-52 (1-275) x US-09-016-366A-22 (1-1081)
 QY 9 LeuProValLeuAlaSerArgAlaIleValAlaProAlaProValGlnAlaLeuGlnGln 28
 Db 2 CTGCCCCGTCCGCGCAGCGCCCTTACCGCGGCTGGCCCAAGCCAGCCGCTCGAGCGCA 61
 QY 29 AlaGlyIleValGlyGlyGlnGlnAlaProArgSerIleTrpProTrpGlnValSerLeu 48
 Db 62 GTGGGCACTGTTGGGGGTTCAGGAGGCCCCCAAGAGCAAGTGGCCCTGGCAAGTGAAGCTG 121
 QY 49 ArgValIleArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
 Db 122 AGAGTCCCGCAGCGATCTGATGATGCACTTCTGGGGGGCTCCCTCATCCACCCCAAGTGG 181
 QY 69 ValLeuThrAlaAlahisCysLeuGlyProAspValIlyAspIleValAlaThrLeuArgVal 88
 Db 182 GTGCTGACCGCGCAGCTGCGTGGGAGCCGAGCTGAAAGATGTGGCCGCGCTCAAGGGTG 241
 QY 89 GlnLeuArgGlnGlnHisIleuTyrTrhisIleuGlnLeuLeuProValSerArgIleIle 108
 Db 242 CAACTGCCGGGAGCAGCACTCTACTACAGGACCAAGCTGCTGCCGCTCAGCAGATCATTC 301
 QY 109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGlnLeuGlu 128
 Db 302 GTGCAACCCACAGTTCTACACCGCCAGATCCGAGGCGGACATTCGCTGTGGAGCTGGAG 361
 QY 129 GlnProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
 Db 362 GAGCGGTGAGAGGTTCCAGCCACGTCACACCGGTCAACCCCTGCTCGAGAGCC 421
 QY 149 PheProProGlyMetProCysTrpValIleThrGlyTrpGlyAspValAspAsnAspGluPro 168
 Db 422 TTCCCGCCCGGGGATGCCGTGGTGCCTGCTGCTGGGGCGGATGTGAAAGTGAAGGAAACCAACTTTGT 481
 QY 169 LeuProProProPheProLeuIlyGlnValIlyValProIleMetGlnAsnHisIleCys 188
 Db 482 CTCCCAACCGCATTTCTCTGAAAGCGGTGAAAGTCCCCCAATAATGAAACCAACTTTGT 541
 QY 189 APVALALVSTRHISLEUVALALATVTRHISGLYASPAASNPVALARGILEILEARXASPAAP 208
 Db 542 GACGGAATAATACCACTTTGGCGCTTACCGGAGAGCGAGTCCGATCTCGGAGCAGAC 601
 QY 209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysIlysGlyAspSerGlyIlyProLeu 228
 Db 602 ATGCTGTGTGGGGGAAACACCCGAGGAGTCAITGCTCCAGGGCGACTCCGGAGGGCCCTTG 661

QY 229 ValCysIysValAaNGlyThrTrpLeuGlnAlaGlyValIleSerTrpAspGluGlyCys 248
 DB 662 GGTGCAAGGTAAGTGGCACTGGCTGGAGCGGGGCTGGTGGAGGCGAGGCTGT 721
 QY 249 AlaGlnProAsnArgProGlyIleTyrThrArgValIleTyrTyrIleuAspTrpIleHis 268
 DB 722 GCCCAGCCCAACCGGCTGGCATCAACCCGCTGCACCTTACTTGGACCTGCATCCAC 781
 QY 269 HisTyrValProIlyValysPro 275
 DB 782 CACTATGTCCCAAAAAGCCG 802

RESULT 8

US-08-978-404B-17
 ; Sequence 17, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard J.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacke, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1081 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-978-404B-17
 Alignment Scores:
 Pred. No.: 1.35e-143 Length: 1081
 Score: 1394.00 Matches: 249
 Percent Similarity: 95.888 Conservatvie: 7
 Best Local Similarity: 93.264 Mismatches: 11
 Query Match: 92.208 Indels: 0
 DB: 2 Gaps: 0
 US-09-598-982C-52 (1-275) x US-08-978-404B-17 (1-1081)
 QY 9 LeuproValIleuAlaSerArgAlaTyrAlaAlaProValIleuGlnGlnGln 28
 DB 2 GTGCCCGTCTGGCGAGCCGCGCTAGCGGGCCCTCCAGGCCAAGCCCTGGAGCGA 61
 QY 29 AlaGlyIleValaGlyGlyGlnGlnAlaProArgSerIleTyrProTrpGlnValSerLeu 48
 DB 62 GTGGGCGATGCTGGGGGTCAAGAGGCGCCCAAGAGCAAGTGGCCCTGGCGAGGCTGG 121

QY 49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerIleuIleHisProGlnTrp 68
 DB 122 AAGATCCCGCCAGCATCTGGATGACCTTGTCCGGGGGCTCCCTCAATCCACCCCAAGTGG 181
 QY 69 ValIleuThrAlaAlaHisCysIleuGlyProAspValIlyAspLeuAlaThrLeuArgVal 88
 DB 182 GTGCTGACCGGAGGCACTGGCTGGAGCCGGAAGCTCAAGAGATCTGGCCCGCCCTCAGGGGTG 241
 QY 89 GlnLeuArgGluGlnHisIleuTyrTyrGlnAspGlnLeuProValSerArgIle 108
 DB 242 CAACGGGGGAGCAGCACTCTACTACAGGACCAAGGCTGGCCGCTCAGAGATCAATC 301
 QY 109 ValHisProGlnPheTyrIleIleGlnThrGlyValAlaAspIleAlaIleuGluLeuGlu 128
 DB 302 GTGCACCCAGCTTCTACACCCGCCAGATCGAGCGGACATCGCCCTGTGGAGCTGGAG 361
 QY 129 GluProValAenIleSerSerArgValHisThrValMetLeuProAlaSerGluThr 148
 DB 362 GAGCCGGTGAAGTCTCCAGCCAGTCCACAGCGGTCAACCTGCCCTCAGAGACC 421
 QY 149 PheProProGlyMetProCysTyrValIleThrGlyTyrGlyAspValAspAspGluPro 168
 DB 422 TTCCCGCCGGGGATGCGGTGGTCACTGGCTGGGGCGATGTGGAATGATGAGCGC 481
 QY 169 LeuProProProPheProLeuIlyValIlyValProIleMetGluAspHisIleCys 188
 DB 482 CTCCACCCAGCATTTCTGTGAAAGCAGGTGAAGGTCCCAATGAAGAAACCAATTTGT 541
 QY 189 AspAlaIlyTyrHisIleuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
 DB 542 GAGCCAAATTCACACTTGGCGCTTACAGCGGAGACAGACGTCGCCGATGTCCGAGAGAC 601
 QY 209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysIlyGlyAspSerGlyIleProLeu 228
 DB 602 ATGCTGTGTGCGGGAAACACCCGGAGGACATCAATGCCAGCGGCGGCGCCCTGTG 661
 QY 229 ValCysIysValAaNGlyThrTrpLeuGlnAlaGlyValIleSerTrpAspGluGlyCys 248
 DB 662 GGTGCAAGGTAAGTGGCACTGGCTGGAGCGGGGCTGGTGGAGGCGAGGCTGT 721

RESULT 9

US-09-917-254-50
 ; Sequence 50, Application US/09917254
 ; Patent No. 6703204
 ; GENERAL INFORMATION:
 ; APPLICANT: Mutter, George
 ; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
 ; FILE REFERENCE: B0801/7224 (URV)
 ; CURRENT APPLICATION NUMBER: US/09/917,254
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIORITY APPLICATION NUMBER: US 60/222,093
 ; PRIORITY FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 50
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-09-917-254-50
 Alignment Scores:
 Pred. No.: 1.35e-143 Length: 1081
 Score: 1394.00 Matches: 249

QY 151 ProGIYMeTProCYeTTPVAlThrGIYTPGIYVAspVAlAspAsnGluPProleuPro 170
 DB 361 CCGGGGATGCGCGTGGCTCACTGGCTGGGCGAGTGGACAAATGATGAGCGCCCTCCCA 420
 QY 171 ProProPheProLeuLYeGInValLYsValProIlleMeTGIUAsnHISleCYAspAla 190
 DB 421 CCGCCATTTCTCTGAAAGAGGTGAAGGTCCCAATTAATGAAACCAACATTTGTGACGA 480
 QY 191 LysTYrHISleuGIYAlATYrThrGIYAspAspValArgIleIeArgAspAspMetleu 210
 DB 481 AAATACACCTTGGCGCTTACACGGGAGACAGCTCCGATGCTCCGTACCACTAGCTG 540
 QY 211 CysAlaGIYAsnSerGIYAspAspSerCysLYsGIYAspSerGIYGIYProleuValCys 230
 DB 541 TGTGCGGGAAACACCCGGAGGACTGATGCTCCAGGGGACTCCGGAGGGCCCTGGTGTGC 600
 QY 231 LysValAenGIYThrTrPLeuGIYAlaGIYValIseTTPAspGIUglYCySAlaGIn 250
 DB 601 AAGGTGAATGGACCTGGCTGACAGGGCGCTGCTGACGTGGGGGAGGGCTGTGCCCA 660
 QY 251 ProAenArgProGIYIleTYrThrArgValIThrTYrLeuAspTTPIlleHISHSIYr 270
 DB 661 CCAACCGGCTGGCATTAACACCGGTGTCACTACTGACTGACTGATCCACCACTAT 720
 QY 271 ValProLYsLYsPro 275
 DB 721 GTCCCAAAAAGCCG 735

RESULT 11
 US-09-079-970A-4
 Sequence 4, Application US/09079970A
 Patent No. 6274366
 GENERAL INFORMATION:
 APPLICANT: Mafiltc, Mark A.
 APPLICANT: Niles, Andrew L.
 APPLICANT: Haak-Frendscho, Mary
 TITLE OR INVENTION: Enzymatically-Active Recombinant Human
 TITLE OR INVENTION: Beta-Trypsin and Method of Making Same
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Intellectual Property Department
 STREET: 8000 Excelsior Drive, Suite 401
 CITY: Madison
 STATE: WISCONSIN
 COUNTRY: U.S.A.
 ZIP: 53717-1914
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/079,970A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Leone, Joseph T.
 REGISTRATION NUMBER: 37,170
 REFERENCE/DOCKET NUMBER: 34506, 073
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 831-2100
 TELEFAX: (608) 831-2106
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 771 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:

ORGANISM: Homo sapiens
 FEATURE: CDS
 NAME/KEY: CDS
 LOCATION: 7..753
 FEATURE: misc signal
 NAME/KEY: misc signal
 LOCATION: 7..18
 US-09-079-970A-4

Alignment Scores:
 Pred. No.: 1.03e-131 Length: 771
 Score: 1284.00 Matches: 227
 Percent Similarity: 95.10% Conservative: 6
 Best Local Similarity: 92.65% Mismatches: 12
 Query Match: 84.92% Indels: 0
 Ds: 3 Gaps: 0

US-09-598-982C-52 (1-275) x US-09-079-970A-4 (1-771)

QY 31 ILeValGIYGIYGIYInGIYAlaProArgSerLYsTPProTTPGIYValIseTTPAspVal 50
 DB 19 ATCGTGGGGGTCAAGAGGCCCAAGAGCAATGGCCCTGGCAAGTGAAGCTTGGATG 78
 QY 51 ArgAspArgTYrTTPMeVHISpHeCYeGIYLYseIleuIleHISpROGIYTPValIleu 70
 DB 79 CACGGCCCATATCTGGATGCACTTCTGGGGGGGCTCCCTCATCCACCCCAAGTGGCTG 138
 QY 71 ThrAlaAlaHISYsLYeGIYProAspValLYsAspLeuAlAThrLeuArgValIleu 90
 DB 139 ACCGACGGCAGCTGGCTGGAGCCGGACCTCAAGATCTGGCCCGCCCTCAAGGATGCACTG 198
 QY 91 ArgGIUglNHISleuTYrTYrGIYAspGIYIleuEupProValIseTTPArgIleIleValHIS 110
 DB 199 CCGAGCAGCACTCTACTACCAAGACAGCTGTGGCTGGCAAGATCATCTGTGCAC 258
 QY 111 ProGIYpHeTYrIleIleGIYThrGIYAlaAspIleAlaLeuGIUglYProleu 130
 DB 259 CCAAGTTCYACCCGCCCAAGATGGAGGGAGCAATTCGCCCTGGAGACTGGAGAGCG 318
 QY 131 ValAsnIleSerSerArgValHISThrValMetLeuProProAlaSerGIUThrPhePro 150
 DB 319 GTGAAGGCTCCAGCAGCTCCACACCGGTCACTCCCTGGCCCTCGACTGACACCTTCCC 378
 QY 151 ProGIYMeTProCYeTTPVAlThrGIYTPGIYVAspVAlAspAsnGluPProleuPro 170
 DB 379 CCGGGGATGCGCGTGGCTCACTGGCTGGGCGAGTGGACAAATGATGAGCGCCCTCCCA 438
 QY 171 ProProPheProLeuLYeGInValLYsValProIlleMeTGIUAsnHISleCYAspAla 190
 DB 439 CCGCCATTTCTCTGAAAGAGGTGAAGGTCCCAATTAATGAAACCAACATTTGTGACGA 498
 QY 191 LysTYrHISleuGIYAlATYrThrGIYAspAspValArgIleIeArgAspAspMetleu 210
 DB 499 AAATACACCTTGGCGCTTACACGGGAGACAGCTCCGATGCTCCGTACCACTAGCTG 558
 QY 211 CysAlaGIYAsnSerGIYAspAspSerCysLYsGIYAspSerGIYGIYProleuValCys 230
 DB 559 TGTGCGGGAAACACCCGGAGGACTGATGCTCCAGGGGACTCCGGAGGGCCCTGGTGTGC 618
 QY 231 LysValAenGIYThrTrPLeuGIYAlaGIYValIseTTPAspGIUglYCySAlaGIn 250
 DB 619 AAGGTGAATGGACCTGGCTGACAGGGCGCTGCTGACGTGGGGGAGGGCTGTGCCCA 678
 QY 251 ProAenArgProGIYIleTYrThrArgValIThrTYrLeuAspTTPIlleHISHSIYr 270
 DB 679 CCAACCGGCTGGCATTAACACCGGTGTCACTACTGACTGACTGATCCACCACTAT 738
 QY 271 ValProLYsLYsPro 275
 DB 739 GTCCCAAAAAGCCG 753

RESULT 12
 US-08-978-404B-1

```

; Sequence 1, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7090
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1031 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; Alignment Scores:
; Pred. No.: 5.09e-115 Length: 1031
; Score: 1134.00 Matches: 23
; Percent Similarity: 82.78% Conservative: 23
; Best Local Similarity: 74.36% Mismatches: 45
; Query Match: 75.00% Indels: 2
; DB: Gaps: 1
;
; US-09-598-982c-52 (1-275) x US-08-978-404B-1 (1-1031)

```

```

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaIrrAlaIlePro 20
DB 27 ATCTTAAAGCTGTGCTGCTCAAGCTGCCCCCTCTCTGCTCCAGCTGGTGAAGCC 86
QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGlnAlaProAr 40
DB 87 GGTCCA-----GCTATMGACAGCAAGAAAGGATTTGTGGGGGAGCAGGAGGACATGGGAAC 140
QY 41 LysTTPProTTPGlnValSerLeuAArgValArgAspArgTTPTrpMetHisPheCysGly 60
DB 141 AAGTGGCCCTCGGAGGAGGAGCCCTGCGCCAAATGACACCTCACTGATGCAATTTCTCGGGT 200
QY 61 GlySerLeuIleHisProGlnIrrTPValIleuThrAlaAlaHisCysLeuGlyProAspVal 80
DB 201 GGCTCCCTCATCCACCAACCAAGTGGTGTCTACTGCGCACTGTGTGGGACCGGATGT 260
QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnHisIleuTyrTyrGlnAspGln 100
DB 261 GCTGACCCCAACAAGGTCAAGATGACGCTCCGTAAGAGATCACTCAATTAACATGACCAAC 320
QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120

```

```

DB 321 CTGATGACTGTGAGCCAGATTCATCAACACCCCGACTTCAATCATGTCAGAGTGGGCA 380
QY 121 AspIleAlaLeuLeuGluLeuGluProValIleAsnIleSerSerArgValHisThrVal 140
DB 381 GACATTTGCCCTCTGAAACTCAAAACCTGTGAACTTTCTGACTAATGTCCACCCCTGTC 440
QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTTPValIrrGlyTTP 160
DB 441 CCCTAACCTCGCTCCAGAGACTTCCCTCAGGAAAGTGTGTGGGTGACAGGCTGG 500
QY 161 GlyAspValAspAsnAspGluProLeuProPheProLeuLeuGlnValIlyVal 180
DB 501 GGTAAACATGACAAATGTGTAAACCTGCCCACTTCTTAAAGAGGTGCAAGTT 560
QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleuGlyAlaTyrThrGlyAsp 200
DB 561 CCCATTATPAAACCAACCTTTGTGTGACTTGAAGTATCAAAAGGCTCATCATCAAGTAC 620
QY 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
DB 621 AATGTCCACATTTGCTCCAGATGACATGCTGTGTGGGAATGAAGACATGACTCCCTGC 680
QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValIleAsnGlyThrTTPLeuGlnAlaGly 240
DB 681 CAGGGGACTCCGAGGACCTCTGTGTGCAAGGATGAGACACCTGCTGACAGGCAAGGC 740
QY 241 ValValSerTTPAspGlyGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
DB 741 GTGTGACAGCTGGGGGAGGGGCTGTGTGACAGCCCAACAGGCTCGACTTACACCCGGGTC 800
QY 261 ThrTyrTyrLeuAspTTPIleHisIleTyrValProLys 273
DB 801 ACCATAACTTGAAGTCCAGATCCACCACATATGCCCCAAG 839

```

```

RESULT 13
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:

```


Db 286 CAGCTTTCGGGGTGCAGCTTCTGTGAGCAGTACTACTAATGGGGACCAAGCTCCTCTCT 345
 QY 104 ValSerArgIleIleValHisProGlnPheTYRlleIleGlnThrGlyAlaSerIleAla 123
 Db 346 TTGAACCGGATCGTGGGACCCCACTATTACAGCGCGGGGGGAGGAGCGTTGGCC 405
 QY 124 LeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeuPro 143
 Db 406 CTGCTGAGGCTTGAAGTCTCCGAAATGCTCCACCAATATCCACCCCATCTCCGCCC 465
 QY 144 ProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspVal 163
 Db 466 CTGCTCCGAGACCTTCCCTCCGAGACATCTGCTGGGTGACAGGCTGGGGCCGACATT 525
 QY 164 AspAsnAspGluProLeuProProPheProPheProLeuIleGlnValLysValProIleMet 183
 Db 526 GATPATATGACGAGCCTCCCACTCCTTATCTCTGATAGCAAGTGAAGGTTCCATTGG 585
 QY 184 GluAsnHisIleLeuAspAlaLysTYRHisLeuGlyAlaTYRThrGlyAspAspValArg 203
 Db 586 GAAAACAGCCTGTGTGACCGGAAAGTACCACTGGGCTTACACGGGAGATGATTTTCCC 645
 QY 204 IleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnhrAspSerCysLysGlyAsp 223
 Db 646 ATTGTCCATGATGGCATCTGTCTGTGAAATACAGAGAGACTCTCCGACAGGCGCAT 705
 QY 224 SerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValAlaSer 243
 Db 706 TCAAGGGGGGCGCATGGTCTGCAAAAGTGAAGGGTAACTGGGCTGCAAGGAGGTCACGC 765
 QY 244 TrpAspGluGlyCysAlaGlnProAsnArgProGlyTlIeTYRThrArgValIThrTYR 263
 Db 766 TGGGGTGAAGGGCTGCCACAGCCCAAGCCTGGGATTCACACCCGGGTGACATCTAC 825
 QY 264 LeuAspTrpIleHisIshisTYRValProLys 273
 Db 826 TTRGACTGGATTCACCGCCTATGTCTCTGAG 855
 RESULT 15
 US-08-845-998-3 / Sequence 3, Application US/08845998
 / Patent No. 5879892
 / GENERAL INFORMATION:
 / APPLICANT: Van Baren, Nicolas
 / APPLICANT: Coullie, Pierre G.
 / APPLICANT: De Smet, Charles
 / APPLICANT: Lucas, Sophie
 / APPLICANT: Boon, Thierry
 / TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
 / NUMBER OF SEQUENCES: 16
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Wolf, Greenfield & Sacks, P. C.
 / STREET: 600 Atlantic Avenue
 / CITY: Boston
 / STATE: MA
 / COUNTRY: US
 / ZIP: 02210
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/845,998
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Van Amsterdarn, John R.
 / REGISTRATION NUMBER: 40,212
 / REFERENCE/DOCKET NUMBER: L0461/7008
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (617)720-3500
 / TELEFAX: (617)720-2441

/ INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 2259 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: double
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA
 / HYPOTHEICAL: NO
 / ANTI-SENSE: NO
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 8..577
 / US-08-845-998-3
 Alignment Scores:
 Pred. No.: 8.51e-113 Length: 2259
 Score: 118.50 Matches: 236
 Percent Similarity: 41.68% Conservative: 17
 Best Local Similarity: 38.88% Mismatches: 22
 Query Match: 73.97% Indels: 333
 DB: 2 Gaps: 2
 US-09-598-982C-52 (1-275) x US-08-845-998-3 (1-2259)
 QY 1 MetLeuSerLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTYRAlaAlaLPro 20
 Db 8 AAGCTGAGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
 Db 68 GCCCAGGCGCAGGCGCTTCAGCAAAAGGCGCATTTGGGGGGGCGAGAGGCCCCCAAGAGC 127
 QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTYRTrpMetHisPheCysGly 60
 Db 128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGGCCCATCTGGATCCACTTCTGCGGG 187
 QY 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 Db 188 GGCCTCCCTCATCCACCACCAGTGGGGTGTAAACCGGGGCGCATGCGTGGAAACCGGACATC 247
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisIleuTYRTrpGlnAspGln 100
 Db 248 AAGGATCTGGCCCGCTCAGGTGCACTCGGGGAGCAGCACTTACTACCAAGGACCAAG 307
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTYRlleIleGlnThrGlyAla 120
 Db 308 CTGCTGCGGCTGACGAGGATCATCGTGCACCCACAGTTCATCATCATCAGACCGGGGCG 367
 QY 121 AsnIleAlaLeuLeuGluLeuGlnubProValAsnIleSerSerArgValHisThrVal 140
 Db 368 GACATCGCCCTGTGAGTGAAGGAGGCCCTGAAACATCTCCAGCAACATCCACACCGGTC 427
 QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTYR 160
 Db 428 AAGCTGCCCGCTGCTCGGAGACCTTCCCGGGGAGTCCGCTGTGGGTCACTGGCTGG 487
 QY 161 GlyAspValAspAsnAsp----- 166
 Db 488 GCGCAGCGTGCACAATAA-TGGTGGGTGTGGGGGACAGCGGGAGCGCGGCGCAGGTGGCA 546
 QY 166 ----- 166
 Db 547 CCAGTCAAGCCACAGGCCAGTCCGTTG99GTTGACAGGGTCCCTGAGGGCGGCTCAGGGA 606
 QY 166 ----- 166
 Db 607 GGGGGACTGTGAGGCCAGATGATGAGAGAGGGGTTGGGAGAGGAGAGGAGAGGAGTGCCT 666
 QY 166 ----- 166
 Db 667 GAGCAGAGCGGTGAGTCCAAAGGCGCTGGGCGTCCCAACCCAGGGGTTTGGAGAGTCC 726
 QY 166 ----- 166

DB 727 CCTTAGCACTCCGATGCTCGGTTTCCCTTCCCTGAAAAGGATGATCAAAAATTGATGAC 786
 QY 166 ----- 166
 DB 787 GTACAGCACTTGTATGTGAGAGAGAAATCAACGCGGGGTCTGTGAAAGGAGAGAGA 846
 QY 166 ----- 166
 DB 847 CCGGTCTGGGATGAGACCTGCGCTCCATCCCTGTGCTACAGACAAGGCAGGGGCC 906
 QY 166 ----- 166
 DB 907 TGGGAATGGGGTCCGTGGGAGTGTGTGGGGGGCTGGAGAACTCACTGTGGCCCTCCA 966
 QY 166 ----- 166
 DB 967 CGAGGACATTTTCACTTAGAAGGTCTGTCCCATTTTATCCACAATTGAGACA 1026
 QY 166 ----- 166
 DB 1027 GCTTTGGGGTACAGCCTGACGCAACCTGGGCTGTGACCTTGGGTCACTCCAGAAAGGAG 1086
 QY 166 ----- 166
 DB 1087 CCTGAGCCACTGTCCCGCTATTCCGCCCAACAAGGGGAACTGAGCCCAAGCCCTG 1146
 QY 166 ----- 166
 DB 1147 TGTTCCTCGGCTGAGGGCAACCGTGAACATGGGCTTAGGCCCAAGAAAGTCAAGCTG 1206
 QY 167 ----- 168
 DB 1207 AGCCGAGGGGAGAGACAGGGTCTGCAACCCCGTGCATGAGCCAGCTTGGCA 1266
 QY 168 ----- 168
 DB 1267 ACCTCAAGGGCCCTCCCTCCCTTCCCAAGTGGGCTTAAATGAGGCCAGGACCCAGG 1326
 QY 168 ----- 168
 DB 1327 ACCAGCCTCAGCGGAGGGGCTGACTGATTCACCGCCCTTCCCGGGGCTGAGAGCA 1386
 QY 168 ----- 168
 DB 1387 CAGAACAGCACTGGGCCCATGTGGCATCTCCCTGCCCCGTGACTCTGCCAACAGTCCA 1446
 QY 168 ----- 168
 DB 1447 CGAAGCAGCACCAGCCGCCCCAAGACCCGGCTCCACGCCCTCCCGCCCCCAGTGCAC 1506
 QY 169 LeuProProPhePheProLeuIySgIValIySValProIleMeGIuAAsnHisIleCys 188
 DB 1507 CTGCCCGCCGCAATACCCGCTGAAGGAGGTGGAAGTCCCGTAGTGGAAAACCACTTTGC 1566
 QY 189 AspAlaIySValIySValIySValIySValIySValIySValIySValIySValIySVal 208
 DB 1567 AACGGGGAATATCACACCCGCTCCATACGGGGCCACAGCTTCAAAATCCTCCGCGATGAC 1626
 QY 209 MetLeuCyValIaGlyAsnSerGlnArgAspSerCyValySgIyAspSerGlyIyProLeu 228
 DB 1627 ATGCTGTGTGGGGGAGCAAAAATCACGACTCTCTGGCAGGGTGACTGTGGAGGGCCCTG 1686
 QY 229 ValCyIySValIaAsnGlyThrTrpLeuGlnaIaGlyValIaSerTrpAspGluGlyCys 248
 DB 1687 GTCTGAAGATGATGGCACTTAACCTGCAAGGGCGCGGTGTCAAGCTGGAGAGAGCTGT 1746
 QY 249 AlaGlnProAsnArgProGlyIleTyThrArgValThrTyTrpLeuAspTrpIleHis 268
 DB 1747 GCCCAGCCCAACCGGCTGGCACTACACCCGTTCACTACTTGGACTGATCCAC 1806
 QY 269 HisTyValProIySValySPro 275

DB 1807 CACTATGTCCCAAGAGGCC 1827

Search completed: August 27, 2005, 23:44:31
Job time : 194 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 16:09:21 ; Search time 6409 Seconds
(without alignments)
254.007 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1512
Sequence: 1 MSLULLLALPVLASRAVAAP.....IYTRVTVYLDVHMHVPRKP 275

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dgapop 6.0 , Dgapext 7.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEVS=xlp
-O=/cgn2.1/USPTO_spool_p/US09598982/rnatc_29082005_080947_26125/app_query.fasta_1.455.
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMax=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09598982@cgn1_1_708@runat_29082005_080947_26125 -NCU=6 -ICPU=3
-NO_WMAP -LARGEOBERRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMOUT=120 -WARN TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002as:
7: geneseqn2002bs:
8: geneseqn2003as:
9: geneseqn2003bs:
10: geneseqn2003cs:
11: geneseqn2003ds:
12: geneseqn2004as:
13: geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1512	100.0	1154	2	AAV44328 Human mas
2	1512	100.0	1154	2	AAV42710 Human mas
3	1512	100.0	1154	3	AAA34959 Human ade
4	1512	100.0	1154	3	AAF21081 Human low
5	1512	100.0	1154	10	AB296775 Human nuc

ID	AAV44328	standard; DNA; 1154 BP.	ADK52587	Human hem	
6	1512	100.0	1154	10	ADK52587 Human hem
7	1512	100.0	1154	11	ABD20624 Human pul
8	1512	100.0	1158	9	ACH04035 Human GDN
9	1512	100.0	1158	12	ADP10384 Reference
10	1512	100.0	1158	13	ADP56069 Human PRO
11	1512	100.0	11713	3	AAA34962 Human ade
12	1512	100.0	17133	3	AAF21084 Human low
13	1512	100.0	17133	10	AB296778 Human nuc
14	1512	100.0	17133	11	ABD20627 Human pul
15	1430	94.6	1142	13	AD085402 Human tum
16	1416	93.7	1334	12	ADQ23979 Human sof
17	1405	92.9	1143	3	AAA34955 Human ade
18	1405	92.9	1143	3	AAF21077 Human low
19	1405	92.9	1143	10	AB296771 Human nuc
20	1405	92.9	1143	10	ADK52589 Hematolog
21	1405	92.9	1143	11	ABD20620 Human pul
22	1405	92.9	1143	12	ADQ6120 Human pul
23	1405	92.9	1143	12	ADQ6120 Human oli
24	1405	92.9	1145	3	AAA34956 Human ade
25	1405	92.9	1145	3	AAF21078 Human low
26	1405	92.9	1145	10	AB296772 Human nuc
27	1405	92.9	1145	10	ADJ34865 Human nuc
28	1405	92.9	1145	11	ABD20621 Human pul
29	1405	92.9	42587	12	ADV61642 Concateme
30	1402	92.7	1137	2	AAV44329 Human mas
31	1402	92.7	1137	2	AAV42711 Human nuc
32	1402	92.7	1137	3	AAA34957 Human ade
33	1402	92.7	1137	3	AAF21079 Human low
34	1402	92.7	1137	10	AB296773 Human nuc
35	1402	92.7	1137	11	ABD20622 Human pul
36	1401	92.7	2662	6	AAD23854 Human pro
37	1400	92.6	1128	2	AAV44330 Human mas
38	1400	92.6	1128	2	AAV42712 Human mas
39	1394	92.2	1081	2	AAV44331 Human mas
40	1394	92.2	1081	2	AAV42713 Human mas
41	1394	92.2	1081	3	AAA34960 Human ade
42	1394	92.2	1081	3	AAF21082 Human low
43	1394	92.2	1081	6	ABL62377 Colom ade
44	1394	92.2	1081	6	ABL66816 Lung can
45	1394	92.2	1081	6	ABL61818 Colom ade

ALIGNMENTS

RESULT 1	AAV44328	standard; DNA; 1154 BP.	ADK52587	Human hem
AAV44328	AAV44328	standard; DNA; 1154 BP.	ADK52587	Human hem
AC	AAV44328;			
AC				
DT	24-NOV-1998	(first entry)		
DE	Human mast cell tryptase alpha nucleic acid sequence.			
DE	Mast cell tryptase alpha; human; MCP-7; mast cell protease 7; blood clot;			
XX	anticoagulant; myocardial infarction; reocclusion; thromboembolism;			
KW	cerebral embolism; thrombosis; therapy; ss.			
KW				
OS	Homo sapiens.			
XX				
XX				
PH	Key	Location/Qualifiers		
FT	CDS	18..845		
FT		/*tag= a		
XX				
PN	W09824886-A1.			
XX				
PD	11-JUN-1998.			
XX				
PF	25-NOV-1997;	97WO-US021620.		
XX				
PR	04-DEC-1996;	96US-0032354P.		
XX				
XX				
PA	(BGHM) BRIGHAM & WOMENS HOSPITAL.			

XX Stevens RL;
 PI 1998-333308/29.
 XX WPI: 1998-333308/29.
 DR P-PSDB; AAW64237.
 XX New compositions containing trypsinase-7, e.g. mouse mast cell protease-7 -
 PT are used to treat clot formation in e.g. myocardial infarction,
 PT reocclusion following angioplasty or pulmonary thrombo-embolism.
 XX
 PS Disclosure; Page 63; 92pp; English.
 XX
 CC This nucleotide sequence includes a coding region for human mast cell
 CC trypsinase alpha (see AAW64237). The invention provides compositions
 CC comprising an isolated trypsinase-7 that may include chimeric proteins that
 CC contain (a) a human trypsinase for all but the active site region and (b)
 CC the substrate-binding pocket of mouse trypsinase-7 or its homologues (see
 CC AAW64233-36); a method for treating a blood clot by administering a
 CC nucleic acid molecule that codes for a trypsinase-7, or an expression
 CC product, to decrease fibrinogen activity; a nucleic acid encoding a
 CC serine protease (SP); and a method of producing a mature SP by expressing
 CC the inactive zymogen in a host cell, and cleaving the enterokinase
 CC susceptibility domain. The trypsinase-7 polypeptides can be used to treat
 CC disorders mediated by undesirable thrombus clot formation such as
 CC myocardial infarction and reocclusion following angioplasty of blood
 CC clots associated with pulmonary thromboembolism, deep vein thrombosis,
 CC cerebral embolism, renal vein and peripheral arterial thrombosis. They
 CC are also useful for all surgical procedures that require decreased blood
 CC clots
 XX
 XX Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5,18e-129 Length: 1154
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-598-982c-52 (1-275) x AAW44328 (1-1154)

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaIlePro 20
 DB 18 ATGCTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 77
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyTyrLeuValGlyGlnGlnAlaProArgSer 40
 DB 78 GCCCCAGTCCAGAGCCCTGTCAGCAAGGGGTATGTCGGGGGTACAGAGGCCCCCAAGAGC 137
 QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 DB 138 AAGTGGCCCTGGCAGGAGGTGAGCTCGAGAGTCCGCGACCGAATCACTGATGCGGG 197
 QY 61 GlySerLeuLeuHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 DB 198 GGCTCCCTCATCCACCCCGAGTGGGCTGACCCCGCGAGCACTGCCCTGGGACCGAGCTC 257
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisIleuTyrTrpGlnAspGln 100
 DB 258 AAGGATCTGGCCACCCTTCAGGGTGCACCTGCGGAGGAGCACTTACTACTCAGAGCAC 317
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 DB 318 CTGCTCGCAGTCAAGGAGATCATCTGTCACCCACAGTTCAGATCAATCCAGATGGAGGG 377
 QY 121 AspIleAlaLeuLeuGluLeuGluProValAsnIleSerSerArgValHisThrVal 140
 DB 378 GATATGCCCCCTGAGGCTGAGGAGGCCGTGAAATCATCTCCAGCCGCGTCCACAGGTC 437
 QY 141 MetLeuProValSerGlnTrpPheProProGlnMetProCysTrpValThrGlyTrp 160
 DB 438 ATGCTGGCCCTCTGGAGACCTTCCCGGGGATGCGGTGGTCACTGGGCTGG 497

QY 161 GIYAEPVAlAspAsnAspGluProLeuProProPheProLeuIleGlnValIysVal 180
 DB 498 GCGGATGTGGCAAAAGATGAGCCCTCCACCGCATTTCCCTGAGAGAGGTGAAGGTC 557
 QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
 DB 558 CCCAAATATGAAACCAACCATTTGTGACCAAAATATCCACTTGGCGCTTACACGGGAGAC 617
 QY 201 AspValArgIleIleArgAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 618 GACGTCCGATATCTCGTACGACATCTGTTGCGGGGAACGACCAAGAGGACTCTCTGCG 677
 QY 221 LysGlyAspSerGlyGlyProLeuValCysIysValIaemGlyThrTrpLeuGlnAlaGly 240
 DB 678 AAGGGCGACTGTGAGGGGCCCTGGTGTGCAAGGTGAAMGGACCTGGCTACAGGGCGGCG 737
 QY 241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
 DB 738 GTGGTCAGCTGGAGAGGGCTGTGCCCAACCGGCTGGGATCTACACCCGCTGTC 797
 QY 261 ThrTyrTyrLeuAspTrpIleHisIleTyrValProLysLysPro 275
 DB 798 ACCTACTACTTGACTGATCCACCACTATGTCTCCCAAAAAGCCG 842

RESULT 2
 AAW42710
 ID AAW42710 strand: cDNA; 1154 BP.
 XX AC AAW42710;
 XX DT 27-OCT-1998 (first entry)
 XX DE Human mast cell trypsinase alpha encoding cDNA.
 XX KW Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
 KW trypsinase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
 KW antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
 KW hyperproliferative skin disease; peptic ulcer; hypersensitivity; asthma;
 KW inflammatory skin condition; human; mast cell trypsinase alpha; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 18..845
 FT /*tag= a
 FT /product= "Human mast cell trypsinase alpha"
 PN MO9833812-A1.
 XX PD 06-AUG-1998.
 XX PF 30-JAN-1998; 98WO-US001865.
 XX PR 05-FEB-1997; 97US-0037090P.
 XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 XX PI Stevens RL, Huang C;
 XX WPI, 1998-437390/37.
 XX DR P-PSDB; AAW63173.
 XX PT Trypsinase-6 complex inhibitor peptides - used to treat mast cell-mediated
 XX PT inflammatory disorders e.g. asthma.
 XX PS Disclosure; Page 44-45; 69pp; English.
 XX CC This cDNA encodes the human mast cell trypsinase alpha which is a homologue
 CC of the mouse mast cell protease (mMCP-6) zymogen. The invention provides
 CC sequences shown in AAW63160 to AAW63169 that are inhibitors of mMCP-6.
 CC These peptides which are trypsinase-6 complex inhibitors, can be used for
 CC treating a mast cell-mediated inflammatory disorder. The inhibitors can

Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-598-982c-52 (1-275) x AAA34959 (1-1154)
 QY 1 MetLeuSerLeuLeuLeuAlaLeuAlaLeuAlaSerArgAlaTyrAlaAlaPro 20
 Db 18 ATGCTAGAGCTGTGCTGCTGGGGCTGCCCGCTCTGGCCAGCCGGCCCTTAGCGGGCCCT 77
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyLeValGlyGlnGlnAlaProArgSer 40
 Db 78 GCCCCAGTCCAGGCGCTGAGCAAGCGGGGTATCGTGGGGGTCAAGAGGCGCCCAAGAGC 137
 QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 Db 138 AAGTGGCCCTGGAGGAGCTGAGAGTCCGCGACCGAATCTGAGTGCCTTCTCGG 197
 QY 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 Db 198 GGTCTCCCTCATCAACCCCACTGGGTGTGACCCGGCGGCACTGCGGAGCCGAGCGT 257
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnHisLeuTyrTyrGlnAspGln 100
 Db 258 AAGGATCTGGCCAGCCCTCAGGGGTGCAACTGCGGAGCAGACCTCTCACTACCAAGCCAG 317
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnHrgVala 120
 Db 318 CTCTGCTCAATCCAGGATCATCGTGCACCCCAAGTTCATCATCCAGACTGGAGCG 377
 QY 121 AspIleAlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrVal 140
 Db 378 GATATGCGCCCTGTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCCGCTCCACAGCGT 437
 QY 141 MetLeuProProAlaSerGlnUhrPheProGlyMetProCysTrpValThrGlyTyr 160
 Db 438 ATGCTGCGCCCTCCCTCGGAGACTTCCCTCCCGGGGATGCGCTGCTGCTGCTGCTGCTG 497
 QY 161 GlyAspValAspAsnAspGlnUhrLeuProProPheProLeuLysGlnValLysVal 180
 Db 498 GGGGATGGAGCATGTGAGCCCTCCACCCCACTTTCCTCGAAGCGGTAAGGTC 557
 QY 181 ProIleMetGlnAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
 Db 558 CCATATATGAAAAACAATTTGTGAGCGAAATACCACTTGGCCGCTTACAGAGGAGAC 617
 QY 201 AspValAlaArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 Db 618 GAGCTCCCGCATTCCTGTGACGATGCTGTGCGCGGGAACGCAAGAGGGGCTCTCTGC 677
 QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
 Db 678 AAGGGGACTCTTGGAGGGGCCCTGGTGTGCAAGTGAATGGCACTGGCTACAGGGGGG 737
 QY 241 ValValSerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
 Db 738 GTGGTGGCTGGAGCAGGGGCTGTGGCCAGCCCAACCGGCTGGCATATACCCCGTGT 797
 QY 261 ThrTyrTyrLeuAspTrpIleHisIleTyrValProLysLysPro 275
 Db 798 ACCTACTACTGTGAGCTGGATCCACCACTATGTCCCCMAAAGCCG 842

RESULT 4
 AAF21081 standard; DNA; 1154 BP.
 XX AAF21081;
 AC AAF21081;
 AC AAF21081;
 DT 14-MAR-2001 (first entry)
 XX
 DE Human low adenosine antisense oligonucleotide related sequence #2648.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 OS Homo sapiens.
 XX MO200062736-A2.
 PN 26-OCT-2000.
 PD 24-MAR-2000; 2000MO-US008020.
 PF 06-APR-1999; 99US-0127958P.
 PR (UYEC-) UNITV EAST CAROLINA.
 PA (UYEC/) NYCE J W.
 XX Nyce JW;
 PI WPI; 2000-679539/66.
 DR
 XX Low adenosine (A) content antisense oligonucleotides which do not trigger
 PT adenosine receptors during metabolism, useful e.g. for treating cancers
 PT and respiratory obstructions.
 PT
 PS Disclosure; Page 881; 1592pp; English.
 PS
 XX The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
 CC surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergic asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention
 CC
 XX Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5.18e-129 Length: 1154
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

QY	61	GIYSerLeuIIehisProGlnTrpValLeuThrAlaIahisCysLeuGlyProAspVal	80
DB	198	GGCTCCCTCATCCACCCCGAGTGGGTGTGACCGCGCCGACCTGGGAGCCGGGACGTC	257
QY	81	LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuThrTyrGlnAspGln	100
DB	258	AAGGATCTGGCCACCTCCAGGAGGCAACGCGGAGCAAGACCTTACATCCAGGACCGAG	317
QY	101	LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnTrpGlyAla	120
DB	318	CTGCTGCCAGTACAGAGGATCCATCGTGCACCCACATTTCAATCCAGACTGGAGGG	377
QY	121	AspIleAlaLeuLeuGlnLeuGlnGluProValAsnIleSerSerArgValHisThrVal	140
DB	378	GATATGCGCCCTGTGGAGCTGGGAGGCCCGTGAACATCTCCAGCCCGCTCCACACGCGTC	437
QY	141	MetLeuProProAlaSerGlnThrPheProProGlyMetProCysThrValIlePheGlyTrp	160
DB	438	ATGCTGCCCTGCTGCTGGAGACTTCCGCCCGGGAGTGCCTGCTGGCTCACTGGCTGG	497
QY	161	GlyAspValAspAsnAspGluProLeuProProPheProLeuIysGlnValIlyVal	180
DB	498	GGGAGTGGAGCATATGATGAGCCCTCCACCGCCATTTCCCTGAAGCAGGTGAAGGTC	557
QY	181	ProIleMetGlnAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp	200
DB	558	CCCATATATGAAACCAACATTTGTGACGGCAAAATACACCTTGGCCCTTACACGGGAGAC	617
QY	201	AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys	220
DB	618	GAGGTCCGGCATCCGCTGACAGACAGCTGTGTCCCGGAGAACCCAGAGGGGACTCTGG	677
QY	221	LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly	240
DB	678	AAGGGGAGCTCTGGAGGGCCCTCGTGTGCAAGTGAATGGCACTGGGTACAGGGGGGG	737
QY	241	ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal	260
DB	738	GTCGTACAGCTGGAGCCAGGGCTGTGCCAGCCCAACCGGGCTGGCATTCACACCCCGTGC	797
QY	261	ThrTyrTyrLeuAspTrpIleHisIleSlyTyrValProLysIlyAspPro	275
DB	798	ACCTACTACTTGGACTGTGATCCACCATATGTCTCCCAAAAAGCCG	842
RESULT	6		
ADKS2587			
ID	ADK52587	standard; DNA; 1154 BP.	
AC	ADK52587;		
DT	06-MAY-2004	(first entry)	
DE		Hematological disorder associated Gene ID 1847.	
KW		cytostatic; antiarrhythmic; antiviral; virucide; hemostatic; nephrotropic;	
KW		cytostatic; thrombolytic; antiparasitic; gene therapy;	
KW		hematologic disorder; cancer; Sickle Cell Anemia;	
KW		Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;	
KW		Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;	
KW		transfusion reaction; Erythroblastosis; mechanical trauma;	
KW		micro-angiopathic hemolytic anemia; parasite infection; gene; ds.	
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	18..845	
XX		/+tag= a	
PN	MO2003065871-A2.		
XX			
PD	14-AUG-2003.		
XX			
PF	28-JAN-2003; 2003MO-US002484.		

XX	04-FEB-2002;	2002US-0354333P.	
PR	28-FEB-2002;	2002US-0360258P.	
PR	15-MAR-2002;	2002US-0364476P.	
PR	26-APR-2002;	2002US-0375626P.	
PR	06-JUN-2002;	2002US-0386494P.	
PR	24-JUN-2002;	2002US-0390965P.	
PR	28-JUN-2002;	2002US-0392480P.	
PR	03-JUL-2002;	2002US-0394128P.	
PR	31-JUL-2002;	2002US-0399783P.	
PR	13-AUG-2002;	2002US-0403221P.	
PR	30-AUG-2002;	2002US-0407045P.	
PR	25-NOV-2002;	2002US-0429048P.	
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
PI	Carroll JM, Healy A, Welch NS, Kelly LM;		
DR	WPI; 2003-731464/69.		
DR	P-PSDB; ADK52588.		
XX			
PT	Identifying a compound capable of treating a hematologic disorder (e.g.		
PT	anemia or leukemia) comprises assaying the ability of the compound to		
PT	modulate the expression or activity of e.g. 131,148, 199 or 12303		
PT	polypeptide or nucleic acid.		
XX			
PS	Disclosure; SEQ ID NO 45; 232pp; English.		
XX			
CC	The invention relates to a method of identifying a compound capable of		
CC	treating a hematologic disorder comprising assaying the ability of the		
CC	compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,		
CC	194, 14993, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,		
CC	13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic		
CC	acid expression or polypeptide activity, thus, identifying a compound		
CC	capable of treating a hematologic disorder. The methods are useful in		
CC	diagnosing, preventing and treating hematological disorders, such as		
CC	cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,		
CC	Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders		
CC	associated with an increased risk of thrombosis, Herpes, Thalassemia,		
CC	antibody-mediated disorders such as transfusion reactions and		
CC	Erythroblastosis, mechanical trauma to red blood cells such as micro-		
CC	angiopathic hemolytic anemias, infections by parasites or chemical		
CC	injuries. The methods may also be used for identifying compounds that		
CC	modulate hematological disorders. This sequence corresponds to one of the		
CC	genes modulated the compound.		
XX			
SQ	Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	5.18e-129	Length: 1154	
Score:	1512.00	Matches: 275	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	100.00%	Indels: 0	
DB:	10	Gaps: 0	
US-09-598-982C-52 (1-275) x ADK52587 (1-1154)			
QY	1	MetLeuSerLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTyrAlaIlePro	20
DB	18	ATGCTGAGGCTCTGCTGCTGCTGCCCTCGCTGGGAGCCGGCTTACGCGGCCCT	77
QY	21	AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer	40
DB	78	GCCCCAGTCCAGCCCTGTGAGCAAGCGGGTATCGTGGGGGTCAAGAGGCCCCCGAGAGC	137
QY	41	LysTrpProTTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly	60
DB	138	AAGTGGCCCTGGCAGGTGAGCTGAGATGCCCGACCAATATGATGATCACTTGGCGGG	197
QY	61	GIYSerLeuIIehisProGlnTrpValLeuThrAlaIahisCysLeuGlyProAspVal	80
DB	198	GGCTCCCTCATCCACCCCGAGTGGGTGTGACCGCGCCGACCTGGGAGCCGGAGACTTC	257

QY 81 LysAapLeuAlaThrlLeuAargValGlnLeuAargGlnHlslLeuTyrTrpGlnAapGln 100
 Db 258 AAGGATCTGGCCACCTCCAGGGTGGAACTGGCGGAGCAGACCTCTACTACCAAGGACCG 317
 QY 101 LeuLeuProValSerArgIleIleValHlsProGlnPheTyrIleGlnThrGlyAla 120
 Db 318 CTGCTGCCAGTACGAGATCTCTGTGACCCCAAGTTCATCATCTCAAGCTGGAGCG 377
 QY 121 AepIleAlaLeuLeuGluLeuGluGluProValAemIleSerSerArgValHlsThrVal 140
 Db 378 GATATCGCCCTCTGAGCTGGAGAGCCCGTGAACATCTCCAGCGCGCTCCACACGGGTC 437
 QY 141 MetLeuProProAlaSerGlnUhrPheProProGlyYmerProGlyTrrValThrGlyTrp 160
 Db 438 AAGCTGCCCCCTGCTCCGAGACCTTCCCGGGGATGCGCTGTGGCTGCTGGCTGG 497
 QY 161 GIYAAPValAAspAsnAAspGluProLeuProProPheProLeuLysGlnValIlySVal 180
 Db 498 GCGGATGTGGACAAATGATGAGCCCTCCACCGCATTTCCCTGAAGCAGGTGAAGTTC 557
 QY 181 ProIleMetGluAAsnHlsIleCysAAspAlaLysTyrHlsLeuGlyAlaTyrThrGlyAAsp 200
 Db 558 CCGATTAATGGAAACCACTTTGTGACGGAAATACCACTTGGCGCTACACAGGGAGAGC 617
 QY 201 AepValAArgIleIleAArgAAspMetLeuCysAlaGlyAAsnSerGlnAAspSerCys 220
 Db 618 GACGTCGGATCATCCGTACACACATGCTGTGTGGCGGAAACAGCAGGGAGCTCCGTC 677
 QY 221 LysGlyAAspSerGlyGlyProLeuValCysLysValAemGlyThrTrpLeuGlnAlaGly 240
 Db 678 AAGGGCGACTCTGGAGGGCCCTGGTGTGCAAGGTGAATGGAACCTGGCTACAGGGCGGC 737
 QY 241 ValValSerTrpAAspGluGlyCysAlaGlnProAemAArgProGlyTlIeTyrThrAArgVal 260
 Db 738 GTGCTCAGCTGGAGCAGAGGGCTGTGCTCCCAAGCCGCAACCGGCTGAGCTACACCCGCTGTC 797
 QY 261 ThrTyrTyrLeuAAspTrpIleHlsIleHlsTyrValProLysLysPro 275
 Db 798 AACTTACTTGTGACTGTGATCCACCACTAATGTGCTCCCAAAAAGCCCG 842
 RESULT 7
 ID ABD20624 standard; DNA; 1154 BP.
 AC ABD20624;
 DT 29-JUL-2004 (first entry)
 XX
 DE Human pulmonary and inflammatory target DNA #235.
 XX
 KW Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;
 KW respiratory tract inflammation; adenosine sensitivity; lung; cancer;
 KW surfactant depletion; anti-allergic; anti-inflammatory; antiasthmatic;
 KW analgesic; hypotensive; immunosuppressive; cytosstatic; cystic fibrosis;
 KW beta-adrenergic agonist; respiratory disease; pulmonary vasconstriction;
 KW respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;
 KW emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;
 KW pulmonary transplantation rejection; ds.
 XX
 OS Homo sapiens.
 FN WO200285309-A2.
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002W0-US013143.
 XX
 PR 24-APR-2001; 2001US-0286036P.
 XX
 PA (EPIC-) EPIGENESIS PHARM INC.
 PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;

PI Miller S, Tang L, Shahbuddin S;
 XX WPI: 2003-093058/08.
 DR
 XX
 PR Pharmaceutical composition for treating asthma, has antisense
 PR oligonucleotide containing less percentage of adenosine, targeted to
 PR nucleic acids associated with lung airway or lung dysfunction, and
 PR bronchodilating agent.
 XX
 XX Claim 15; SEQ ID NO 12017; 763bp; English.
 CC This invention describes a novel composition (a) a first active agent,
 CC comprising oligonucleotides, effective for alleviating
 CC bronchoconstriction, respiratory tract inflammation, allergies and
 CC reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,
 CC surfactant depletion or hyposecretion, when administered to a mammal. The
 CC oligonucleotides are derived from a gene encoding or regulating
 CC expression of a target polypeptide associated with lung airway or lung
 CC dysfunction or cancer and can be anti-sense to the corresponding mRNA.
 CC The invention also describes a kit, that comprises: (a) a delivery
 CC device, in separate containers, (b) the oligonucleotides, (c)
 CC instructions for adding a carrier and for use of the kit. The composition
 CC of the invention has anti-allergic, anti-inflammatory, antiasthmatic,
 CC analgesic, hypotensive, immunosuppressive and cytosstatic activity, is a
 CC beta-adrenergic agonist. The composition is useful for preventing or
 CC treating a respiratory, lung or malignant disease. The administered
 CC composition comprises oligo and is administered to reduce the production
 CC or availability, or to increase the degradation of the target mRNA or to
 CC reduce the amount of target polypeptide present in the lungs. The
 CC pulmonary obstruction, and/or bronchoconstriction and/or lung
 CC inflammation, allergies and/or surfactant hypoproduction are associated
 CC with a disease or condition such as pulmonary vasconstriction.
 CC Inflammation, allergies, asthma, impeded respiration, respiratory
 CC distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary
 CC transplantation rejection, pulmonary infections, bronchitis or cancer.
 CC The reduced adenosine content of the anti-sense oligos corresponding to
 CC thymidines present in the target RNA serves to prevent the breakdown of
 CC the oligonucleotides into products that fuse adenosine into the system
 CC e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to
 CC prevent any unwanted effects due to it
 XX
 SQ Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.18e-129 Length: 1154
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-598-982C-52 (1-275) x ABD20624 (1-1154)
 QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 Db 18 AAGTGAAGCCCTGCTGCTGCTGCGGTGCGCTCCGCTGGAGCCCGGCTTACGGCCCTT 77
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
 Db 78 GCCCAGTCCAGGGCCCTGCAAGCAGGGGTATGCTCGGGGGTCAAGAGGCCCCAGAGAGC 137
 QY 41 LysTrpProTrpGlnValSerLeuAArgValAArgAAspArgTyrTrpMetHlsPheCysGly 60
 Db 138 AAGTGGCCCTGCGAGGAGGCTGAGGCTCCGACCAAGATCTGTAAGCACTTCTGGCGGG 197
 QY 61 GlySerLeuIleHlsAProGlnTrpValLeuThrAlaHlsCysLeuGlyProAAspVal 80
 Db 198 GACTCCCTCATCCACCCCACTGGAGTGTGACCCCGGGGCACTCCCTGGAGACCGGAGCTC 257
 QY 81 LysAapLeuAlaThrlLeuAargValGlnLeuAargGlnHlslLeuTyrTrpGlnAapGln 100
 Db 258 AAGGATCTGGCCACCTCCAGGGTGGAACTGGCGGAGCAGACCTCTACTACCAAGGACCG 317

QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 DB 318 CTGCTCCAGTCCAGCAGATCATCGTGCACCACCACTTACATCATCCAGCATGGAGCG 377
 QY 121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
 DB 378 GATATGCCCTGTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCTCCACAGCGGTC 437
 QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTyrValThrGlyTyr 160
 DB 438 ATGCTGCCCCCTGCTCCGAGACCTTCCCCCGGGATGCGGTGGGTCACTGGCTGG 497
 QY 161 GlyAspValAspAsnAspGluProLeuProProPheProIleuLysGlnValLysVal 180
 DB 498 GCGGATGTGGACAAATGATGAGCCCTCCCAACCCGCAATTTCCCTGAAAGCAGGAGGTC 557
 QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleuGluValAlaTyrThrArgVal 200
 DB 558 CCCATATGGAAACCAACATTTGTGACGGAAATACCCACTTGGCCCTTACACAGGGAGAC 617
 QY 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 618 GACGTCGGCATCATCCGTGACGACATGCTGTGTGCGGGAAACGACGAGGGACTCTGCG 677
 QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTyrLeuGlnAlaGly 240
 DB 678 AAGGGGACTCTGGAGGGCCCTGCTGTGCAGATGGCACTGGCTACAGCGGGCG 737
 QY 241 ValValSerTyrAspGluGlyCysAlaGlnProAsnArgProGlyTyrIleTyrThrArgVal 260
 DB 738 GTGGTCAAGCTGGAGCAGGGGCTGTGCCCCAGCCCAACCGGGCTGGCATTCACACCCGTC 797
 QY 261 ThrTyrTyrLeuAspTyrIleHisHisTyrValProLysLysPro 275
 DB 798 ACCTACTACTGGACTGGANTCCACCACTATGTCTCCCAAAAAGCGG 842
 RESULT 8
 ACH04035 ID ACH04035 standard; cDNA; 1158 BP.
 XX ACH04035;
 XX
 DT 26-SEP-2003 (first entry)
 DE Human cDNA differentially expressed in lung cancer #240.
 XX
 KW Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
 KW respiratory disorder; lung cancer; asthma; human.
 OS Homo sapiens.
 XX
 PN US2003065157-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 04-APR-2002; 2002US-00116802.
 XX
 PR 04-APR-2001; 2001US-0281593P.
 XX
 PA (LASE/) LASEK A W.
 XX
 PI Lasek AW;
 DR WPI; 2003-540803/51.
 XX
 PT New combination comprising cDNAs that are differentially expressed in
 PT respiratory disorders, useful for diagnosing or treating respiratory
 PT disorders e.g., lung cancer, chronic obstructive pulmonary disease,
 XX emphysema or asthma.
 PS Claim 1; Page: 39pp; English.
 XX

CC The invention relates to a combination comprising cDNAs or their
 CC complements that are differentially expressed in respiratory disorder.
 CC The combination is useful for preparing a composition for diagnosing or
 CC treating respiratory disorders e.g. lung cancer, chronic obstructive
 CC pulmonary disease, emphysema or asthma. The present sequence represents
 CC human cDNA differentially expressed during lung cancer
 XX
 SQ Sequence 1158 BP; 203 A; 413 C; 332 G; 210 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,2e-129 Length: 1158
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 9 Gaps: 0
 US-09-598-982C-52 (1-275) x ACH04035 (1-1158)
 QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaPro 20
 DB 18 ATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 77
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
 DB 78 GCCCAAGTCCAGGCCCTGGAGCCAGGATTCGTCCGGGGGTCAAGAGCCGCCAGAGAC 137
 QY 41 LysTyrProTyrGlnValSerLeuArgValArgAspArgTyrTyrMetHisPheCysGly 60
 DB 138 AAGTGGCCCTGGCAGGTGAGCTGAGAGTCCGCGACCCGATCTGATGCATCTTGGCGGG 197
 QY 61 GlySerLeuIleHisProGlnTyrValLeuThrAlaAlaHisCysValLeuGlyProAspVal 80
 DB 198 GGCCTCCCTCATCCACCCCAAGGGGTGTGACCCGGGGCGCATGCTGGAGACCCGAGCG 257
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisIleuTyrTyrGlnAspGln 100
 DB 258 AAGGATCTGGCCACCTTCAAGGTGCACTGGGGAGCGACACTTACTACAGAGACAG 317
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 DB 318 CTGCTCCAGTCCAGCAGATCATCGTGCACCACCACTTACATCATCCAGCATGGAGCG 377
 QY 121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
 DB 378 GATATGCCCTGTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCTCCACAGCGGTC 437
 QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTyrValThrGlyTyr 160
 DB 438 ATGCTGCCCCCTGCTCCGAGACCTTCCCCCGGGATGCGGTGGGTCACTGGCTGG 497
 QY 161 GlyAspValAspAsnAspGluProLeuProProPheProIleuLysGlnValLysVal 180
 DB 498 GCGGATGTGGACAAATGATGAGCCCTCCCAACCCGCAATTTCCCTGAAAGCAGGAGGTC 557
 QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleuGluValAlaTyrThrArgVal 200
 DB 558 CCCATATGGAAACCAACATTTGTGACGGAAATACCCACTTGGCCCTTACACAGGGAGAC 617
 QY 201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 618 GACGTCGGCATCATCCGTGACGACATGCTGTGTGCGGGAAACGACGAGGGACTCTGCG 677
 QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTyrLeuGlnAlaGly 240
 DB 678 AAGGGGACTCTGGAGGGCCCTGCTGTGCAGATGGCACTGGCTACAGCGGGCG 737
 QY 241 ValValSerTyrAspGluGlyCysAlaGlnProAsnArgProGlyTyrIleTyrThrArgVal 260
 DB 738 GTGGTCAAGCTGGAGCAGGGGCTGTGCCCCAGCCCAACCGGGCTGGCATTCACACCCGTC 797
 QY 261 ThrTyrTyrLeuAspTyrIleHisHisTyrValProLysLysPro 275

DB 798 ACCTACTGTGACTGATCCACCACATGATGTCGCCAAAAGCCG 842

RESULT 9
ADP10384
ID ADP10384 standard; DNA; 1158 BP.

AC ADP10384;
XX ADP10384;
XX 12-AUG-2004 (first entry)

DE Reference mRNA sequences for marker probe #61.

DE transplamt rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.

OS Homo sapiens.

PN WO2004042346-A2.

PD 21-MAY-2004.

PF 24-APR-2003; 2003WO-US012946.

PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.

PA (EXPR-) EXPRESSION DIAGNOSTICS INC.

PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M,
PI Rosenberg S;
DR MPI; 2004-400724/37.

PT Diagnosing or monitoring transplamt rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplamt
PR rejection, in an individual, comprises detecting the expression level of
PR the gene.

PS Claim 80; SEQ ID NO 393; 1762bp; English.

CC The present invention relates to diagnosing or monitoring transplamt
CC rejection, e.g. cardiac or kidney transplamt rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplamt rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplamt rejection,
CC xenotransplamt rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.

SO Sequence 1158 BP; 203 A; 413 C; 332 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.2e-129 Length: 1158
Score: 1512.00 Matches: 275
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Gaps: 0
DB: 12 Gaps: 0

US-09-598-982c-52 (1-275) x ADP10384 (1-1158)

QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValIleuAlaSerAraAlaTrrAlaAlaPro 20
DB 18 ATGCTGAGGCTGCTGCTGGGCGCTGCCCTCCTGCGGCAAGCCGCTCAAGCGGCTCT 77
QY 21 AAlaProValAlaAlaLeuGlnAlaGlyIleValAlaGlyIleGlnAlaProAraGser 40

DB 78 GCCCAGTCCAGGCCCTGACAGCCGAGGATCGTCCGGGGTCCAGAGGCCCCCGAGAGC 137
QY 41 LysTrpProTrrPglValIserLeuAraValAraAraAraGrrYrTrpMetHisAphCyAsGly 60
DB 138 AAGTGGCCCTGAGGTAGGCTGAGAGTCCGCACTGATGAGACCTTCTGCGGG 197
QY 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCySLeuGlyProAraVal 80
DB 198 GGCTCCCTCATCCACCCCACTGGGTGCTGACCCGGGGGCACTGCCCTGGACCCGGAGCTC 257
QY 81 LysAraLeuAlaThrLeuAraValGlnLeuAraGlnGlnHisLeuYrTrpGlnAraGln 100
DB 258 AAGGATCTGGCCACCCCTGAGGTGCACTGGGGAGCGCACCTCTACTACAGAGCCAG 317
QY 101 LeuLeuProValIserAraGlyIleValHisProGlnPheYrTrpIleGlnThrGlyAla 120
DB 318 CTGCTGCAGTCAAGAGGATCATGCTGCAACCACAGTTCATCATCATCAAGCTGAGAGCG 377
QY 121 AspIleAlaLeuLeuGlnLeuGlnIleProValAsnIleSerSerAraValHisThrVal 140
DB 378 GATATCGCCCTGCTGGAGCTGAGAGGCCCTGAAACATCTCCAGCCGCTCCACACGGCTC 437
QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCystrpValThrGlyTrp 160
DB 438 ATGCTGCCCTGCTGGAGACCTTCCCGGGGATGCGCTGCTGGGTCACTGGCTGG 497
QY 161 GlyAraProAlaAraAraAraGlnProLeuProProPheProLeuAraGlnValIleVal 180
DB 498 GGCGATGGGCAATGATGAGCCCTCCACCCGCACTTCCCTGAAACAGGTGAAGGCTC 557
QY 181 ProIleMetGluAraHisIleCySAspAlaIleYrTrpHisLeuGlyAlaIleYrThrGlyAraP 200
DB 558 CCCATTAATGAAAACCACTTTGTGAGCGCAAAATRACACTTGTGGCGCTTACAGCGGAGAC 617
QY 201 AspValAraGlyIleIleAraAraAraAraAraAraAraAraAraAraAraAraAraAraAra 220
DB 618 GACGTCGCGCATCATCCGTCAGACATGCTGTGCTGGGGAACAGCGAGGAGCACTCCGTC 677
QY 221 LysGlyAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAra 240
DB 678 AAGGGCGACTGAGAGGCGCCCTGGTGTGCAAGTGAAGTGGACCTGCTACAGCGGGGAC 737
QY 241 ValValIserTrpAraGlyGlyCySAlaGlnProAraAraAraAraAraAraAraAraAraAra 260
DB 738 GTGCTCAGCTGGAGACGAGGGCTGTGCGCCAGCCCAACCGCCCTGCGATCTACACCCGTCTC 797
QY 261 ThrYrTrpLeuAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAraAra 275
DB 798 ACCTACTGATGACTGATCCACCACTAATGTCGCCAAAAGCCG 842

RESULT 10
ADP56069
ID ADP56069 standard; cDNA; 1158 BP.

AC ADP56069;

DT 18-NOV-2004 (first entry)

DE Human PRO cDNA sequence SEQ ID NO:2045.

DE human; PRO; immune related disease; inflammatory immune response;
DE immune response stimulation; antiinflammatory; antiinflammatory; antiarthritic;
DE antiasthmatic; antidiabetic; antiinflammatory; antiemetic; antiinfectious;
DE antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
DE haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
DE nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
DE viricide; gene therapy; gene; ss.

OS Homo sapiens.
PN WO2004039956-A2.
PD 13-MAY-2004.

XX Nyce JW;
XX MPI: 2000-205971/18.

PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT bronchiectasis, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.

Dislosure; Page 810-814; 1343pp; English.

The present invention describes a new composition comprising an antisense
oligonucleotide (ON) with low adenosine (up to 15%), which targets
nucleic acids involved in bronchoconstriction, allergies, and/or
inflammation. The ON can have anti-inflammatory, anti-allergic,
anti-asthmatic, cytoskeletal and analgesic activities. The compositions are
useful for the treatment of diseases associated with inflammation,
impaird always, including lung disease and diseases whose secondary
effects afflict the lungs of a subject. They can be used for treating
e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
impeded respiration, respiratory distress syndrome, pain, cystic
fibrosis, pulmonary hypertension, emphysema, chronic obstructive
pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
carcinomas, and cancers which may metastasise to the lungs, including
breast and prostate cancer. The reduction of the adenosine content of the
ONs reduces side effects. The A-containing ONs break down with the
release of deoxyadenosine which activates adenosine receptors causing the
bronchoconstriction and inflammation. AAA32313 to AAA5312 represent the
nucleotide sequences given in the sequence listing from the present
invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
sequences are also called SEQ ID NO:1 to 185, but the sequences differ
from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3323 to
AAA3392) are specifically claimed ONs from the present invention. N.B.
Sequences given in the disclosure of the present invention do not match
listing with their corresponding SEQ ID NO: sequences given in the sequence
listing

Sequence 17133 BP: 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.: 1.54e-127 Length: 17133
Score: 1512.00 Matches: 275

Percent Similarity: 100.00% Conservattive: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-598-982C-52 (1-275) x AAA434962 (1-17133)

QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
DB 12636 AAGCTGAGGCTCTGTGTGGCGTCCGCTCCAGGCGGCGCCCTACCCGCGCCCT
QY 21 AlaProValGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
DB 12696 GCCCCAGTCCAGGCGCTGAGCAAGGCGGATATCGTCGGGGGTCAGAGAGCCGCCAGGAC
QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
DB 12756 AAGTGGCCCTGGCAGGTGAGCGCTGAGAGTCCCGCAGCCAGTACTGGATGTCAGCCGGG
QY 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
DB 12816 GGCTCCCTATCCACCCCGAGGGGTGCGACCCGCGCCAGTCCCTGGGACCGGACGTC
QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisIleuTyrTyrGlnAspGln 100
DB 12876 AAGGATCTGGCCACCTCAGGTTGCAACTGCCGGAGCAGCACTTACTACAGAGACAG 12935
QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
DB 12936 CTGGCTGCGAGTCAAGCATTCGTGCGACCCACAGTTTCAATCATCCAGAGCTGGAGGG

QY 121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
DB 12996 GATATCGCCCTGCTGGAGCTGGAGAGGCCCGCTGAACATCTCCAGCGGCTCCACACCGCTC 13095

QY 141 MetLeuProProAlaSerGluTrpPheProProGlyMetProCysTrpValThrGlyTrp 160
DB 13056 ATGCTGCCCTGCTCGAGAACCTTCCCGCGGGAGTCCGCTGCTGGTCACTGGCTGG 13115

QY 161 GlyAspValIaAspAspGluProLeuProProPheProLeuLysGlnValLysVal 180
DB 13116 GGCGATGGAGCAATGAGAGAGCCCTCCACCGCCATTTCCCTGAAAGCAGGTGAAGGTC 13175

QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleuGlyAlaTyrThrGlyAsp 200
DB 13176 CCGATATGAGAAAACCACTTTGTGAGCCGCAAAATRACACTTGGCGCTTACAGGGAGAC 13235

QY 201 AspValArgIleIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
DB 13236 GACGTCGCATCATCCCGTGAAGCAGCATGCTGTGTCGGGAAACAGCCAGAGGAGCTCC 13295

QY 221 LysGlyAspSerGlyGlyProLeuValCysIleValAsnGlyTrpTrpLeuGlnAlaGly 240
DB 13296 AAGGCGCATCTGTGAGAGGCGCCCTGCTGTCGAAAGTGAATGGCACCTGACAGGCGGCG 13355

QY 241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
DB 13356 GTGTGAGCTGGAGAGGCGCTGTGTCGCAAGCCCAACCGGCTTGGCATTTACCGCTTC 13415

QY 261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
DB 13416 ACCATCATCTTGGAGCTGGATCCACCACTATGTCCCAAAAAGCGG 13460

RESULT 12
AAF21084
ID AAF21084 standard; DNA; 17133 BP.
XX AC
XX AAF21084;
XX 14-MAR-2001 (first entry)
DT
XX
XX Human low adenosine antisense oligonucleotide related sequence #2651.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
human, airway disorder; bronchoconstriction; lung inflammation;
surtactant depletion; respiratory; bronchodilator; anti-inflammatory;
immunosuppressive; anti-asthmatic; analgesic; hypotensive; cyotostatic;
respiratory obstruction; pulmonary obstruction; impeded respiration;
surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
pulmonary hypertension; emphysema; pulmonary transplantation rejection;
chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
cancer; ss.
OS Homo sapiens.
XX WO200062736-A2.
XX 26-OCT-2000.
XX 24-MAR-2000; 2000WO-US008020.
XX 06-APR-1999; 99US-0127958P.
XX (UYEC-) UNIV EAST CAROLINA.
XX (UYCE/) NYCE J W.
XX Nyce JW;
XX MPI: 2000-679539/66.
DB 12936 CTGGCTGCGAGTCAAGCATTCGTGCGACCCACAGTTTCAATCATCCAGAGCTGGAGGG
PT Low adenosine (A) content antisense oligonucleotides which do not trigger
PR adenosine receptors during metabolism, useful e.g. for treating cancers

PT and respiratory obstructions.

XX Dislosure; Page 882-887; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotide the A is replaced by a 'universal' or alternative base. (1) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiaesthetic, hypotensive and cytostatic activities. CC The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with CC lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, CC immunoglobulins and antibodies, antibody receptors, cytokines and CC chemokines, endogenously produced specific and non-specific enzymes, CC binding proteins, adhesion molecules and their receptors, cytokine and CC chemokine receptors, adenosine receptors, bradykinin receptors, central CC nervous system (CNS) and peripheral nervous and non-nervous system CC receptors, CNS and peripheral nervous and non-nervous system peptide CC transmitters, defensins, growth factors, vasoactive peptides and CC receptors, binding proteins and malignancy associated proteins. The CC antisense oligonucleotides may be used in this way to treat disorders CC including respiratory obstruction (especially pulmonary obstruction CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or CC surfactant hypoproduction which are associated with a disease or CC condition selected from pulmonary vasoconstriction, inflammation, CC allergic asthma, impaired respiratory, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD), CC pulmonary transplantation rejection, pulmonary infections, bronchitis, CC and/or cancer. AAF18434 to AAF21543 represent human polymucleotide CC fragments and antisense oligonucleotides used in the exemplification of CC the present invention

SO Sequence 17133 BP; 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:

Pred. No.: 1.54e-127 Length: 17133
Score: 1512.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-598-982C-52 (1-275) X AAF21084 (1-17133)

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaIlePro 20
Db 12636 ATGCTGAGCTGCTGCTGGCGGCTGCCGCTGCGGAGCCGCTTACGGGCCCCCT 12695
QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
Db 12696 GCCCCAGTCCAGGCGCTGCGAGCAAGGGGATATCGTGGGGGTGAGGAGGCCCCAGGAGG 12755
QY 41 LysTyrProTyrGlnValSerLeuAlaArgAspArgTyrTyrMetHisPheCysGly 60
Db 12756 AAGTGGCCCTGGGAGGTGAGCCTGAGAGTCCGGGACCGAATCGATGTCCTCTCGGGG 12815
QY 61 GlySerLeuLeuHisProGlnTyrValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
Db 12816 GGGCTCCCTCAATCCACCCCAAGGGGTGCTAACCAGCCGCGGAGCCGAGACGTC 12875
QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisSerValTyrTyrGlnAspGln 100
Db 12876 AAGGATCTGGCCACCTCCAGGGTGCACACTCGGGAGAGAGACCTCTTACACAGACCCAG 12935
QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
Db 12936 CTGCTGCGCAGTCCAGAGGATCATCTGTCACCCACAGTTTCAATCCAGATCGAGAGCG 12995
QY 121 AspTleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
Db 12996 GATATATCCCTGCTGGAGCTGGAGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTCC 13055

QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTyrValThrGlyTyr 160
Db 13056 ATGCTGCCCCCTCCCTCCGAGACCTTCCCCCGGGGAGTCCGCTGGGTGACTGGCTGG 13115

QY 161 GlyAspValAspAsnAspGlnProLeuProProPheProLeuLeuValValValVal 180
Db 13116 GCGGATGTGACAAATGATGAGCCCTCCACCGCATTTCCCTTGAAGCAGGTGAAAGTTC 13175

QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleValGlyAlaTyrThrGlyAsp 200
Db 13176 CCCATTAATGGAAACCAACATTTGTGAGCGAAATATCCACTTGGGCGCTTACACGGGAGAC 13235

QY 201 AspValArgIleIleLysAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
Db 13236 GACGTCGCCGATCATCCGTCAGCACATGCTGTGTCCGGGAAACAGCAAGGAGGACTCC 13295

QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyTyrThrLeuGlnAlaGly 240
Db 13296 AAGGGGACTCTGGAGAGGCCCCCTGGTGTGCAAGGTGAATGGCACCTTGGCTTACAGGCGGG 13355

QY 241 ValValSerTyrAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
Db 13356 GTGTCAGCTGGGAGAGGGGCTGTGCCAGCCCAACCGGCGTGGATCTACACCCGCTGTC 13415

QY 261 ThrTyrTyrLeuAspTyrIleHisHisTyrValProLysLysPro 275
Db 13416 ACCCTACTACTGGACTGGATCCACCACTAATGTCCCCCAAAAAGCCG 13460

RESULT 13
ABZ96778
ID ABZ96778 standard; DNA; 17133 BP.
XX
AC ABZ96778;
XX
DT 17-OCT-2003 (first entry)
XX
DE Human nucleic acid sequence.
XX
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW antiaesthetic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
OS Homo sapiens.
XX
PN WO200285308-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013135.
XX
PK 24-APR-2001; 2001US-0286137P.
XX
PA (EPIG-) EPIGENESIS PHARM INC.
XX
PI Nyce JM, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
DR WPI; 2003-229219/22.
XX
PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX
PS Dislosure; SEQ ID NO 12020; 872pp; English.
XX
CC The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 15:02:36 ; Search time 3552 Seconds
(without alignment) 3751.460 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1512
Sequence: 1 MSHLLMLLPLVLSRAAYAAP.....IYTRVTVYLDWIMHIVPKKP 275

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delp 6.0 , Delpext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spoop_p/US09598982/runat_29082005_080946_26113/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastlap -SUPERX=rgc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=humand0.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODB=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982@cgn_1_1_5600@runat_29082005_080946_26113 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOBJECT -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIDLOC
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELTEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1512	100.0	1154	6	AR080459 Sequence
2	1512	100.0	1154	9	HUMTRY M30038 Human trypt
3	1494	98.8	884	9	AF206665 Homo sapi
4	1494	98.8	1206	9	BC028059 Homo sapi

Result	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN
RSRULT 1	AR080459	Sequence 11 from patent US 5968782.	AR080459	AR080459.1	GI:10007194	Unknown.	Unknown.	1 (bases 1 to 1154)	Stevens, R.L.	Mast cell protease that cleaves fibrinogen	Patent: US 5968782-A 11 19-OCR-1999;	Location/Qualifiers	1..1154 /organism="unknown" /mol_type="unassigned DNA"

ALIGNMENTS

Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
7.85e-136	1512.00	100.00%	100.00%	100.00%	6		1154	275	0	0	0	0

438 ATGCTGCCCTGCTCGGAGACCTTCCCTCCCGGGAGTCGGTGGGTCACTGGCTG 497

161 G1YAAPVAlAspAAspGluProleuProProProPheProleuLysGlnVallyVal 180
 498 GGGGATGTGAGAAATGATGAGCCCTCCAGCCGCAATTTCCCTGAAGGATGAAAGTTC 557

181 ProIleMeTcG1uAAsnHisIleCysAspAlaLysTyrHisIleuG1yAlaTyrThrG1yAsp 200
 558 CCCATTAATGAAAACCACTTTGTGACGCAAAAATACCACTTGGCCCTCAKACGGGAGAC 617

201 AspValArgIleIleArgAspAAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 618 GACGTCGGCATCATCCGTGACGACATGCTGTGTGCGGGGAAACAGCAGGAGGACTCTCC 677

221 LysG1yAspSerG1yG1yProleuValCysLysValAsnG1yThrTrpLeuGlnAlaGly 240
 678 AAGGGGAGCTCTGGAGGGCCCTGTGTCAAGGTAATGCAACCTGGCTAACAGCGGGGC 737

241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProG1yIleTyrThrArgVal 260
 738 GTGGTCAGCTGGAGCGAGGGCTGTGCTCCAGCCCAACCGGCTGGCATCTACACCGGTTC 797

261 ThrTyrTyrLeuAspTrpIleHisTyrValProLysLysPro 275
 798 ACCTACTACTTGAGACTGATCCACCACTATGTCTCCCAAAAAGCCG 842

RESULT 3
 AF206665 884 bp mRNA linear PRI 24-OCT-2002
 LOCUS Homo sapiens mast cell alpha II tryptase mRNA, complete cds,
 DEFINITION alternatively spliced.
 ACCESSION AF206665.1 GI:11493897
 VERSION AF206665.1
 KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 884)
 Wang, H.W., McNeill, H.P., Thomas, P.S., Murphy, B.N., Webster, M.J.,
 Hettiaratchi, A., King, G., Heywood, G.J., Huang, C., Stevens, R.L. and
 Hunt, J.E.
 delta Tryptase Is Expressed in Multiple Human Tissues, and a
 Recombinant Form Has Proteolytic Activity
 J. Immunol. 169 (9), 5145-5152 (2002)
 2 (bases 1 to 884)
 Hunt, J.E., Wang, H.W., Thomas, H.P. and McNeill, H.P.
 Direct Submission
 Submitted (19-NOV-1999) Pathology, University of New South Wales,
 High St., Kensington, NSW 2052, Australia
 Location/Qualifiers
 1. 884
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p13.3"
 /issue_type="Lung"
 /clone_lib="Invitrogen cat. number D6090-15"
 11. 838
 /note="Serine protease; alternatively spliced; similar to
 the product encoded by GenBank Accession Number AF098328"
 /codon_start=1
 /product="mast cell alpha II tryptase"
 /protein_id="AA035695.1"
 /db_xref="GI:11493898"
 /translation="MISLILALPVLASPAAYAAPVQALQAGIVGGGAPRSKMP
 QVSLRVRDRYMMHFCGSLIHPQWVLTAAHCLGSPVXKDLATLRVLRBGLLYQDL
 PYSRIIVHPQFYIQTGADIALIELEPKNVSISSRHVTLWPASETFPDMLCAGNTRD
 GDVNDPEPLPPEPLKQVAVPIMENHIQAKHLGALYTDVRIIRDDMLCAGNTRD
 SCGGDSGGPLVCKVNGTWTQAGVAVSWDECAQPNRPGIYTRVTVYLDIMTHYVPPK"

ORIGIN

Alignment Scores:
 Pred. No.: 3.11e-134 Length: 884
 Score: 1494.00 Matches: 271
 Percent Similarity: 99.64% Conservative: 3
 Best Local Similarity: 98.55% Mismatches: 1
 Query Match: 98.81% Indels: 0
 DB: 9 Gaps: 0

US-09-598-982c-52 (1-275) x AF206665 (1-884)

OY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTyrValAlaPro 20
 DB 11 ATGCTGAGCCCTGCTGCTGCTGCGCTGCGCTGCTGCGGAGCCCGGCTTAAGCGGCCCT 70

OY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyTLeValGlyGlyGlnAlaProArgSer 40
 DB 71 GCCCAGTCCAGGCCCTTCAGCAAGCGGGTATCGTGGGGGTCAAGAGGCCCCCAAGGAGC 130

OY 41 LysTrpProTrpGlnValAsnSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 DB 131 AAGTGGCCCTGGCAGGTGAGGCTGAGAGTCCGCAACGATACGATGACACTTCTGTGGG 190

OY 61 GlySerLeuIleHisProGlnTrpValIleuThrAlaHisCysLeuGlyProAspVal 80
 DB 191 GGCTCCCTCATCCACCCCAAGTGGGTGTGACCCGGGGGCACTGCTGGAGCCGGACGTC 250

OY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGln 100
 DB 251 AAGGATCTGGCCACCCCTCAGGGGTCAACTGGGGAGCAGCACTTACTTACCAAGGACGAG 310

OY 101 LeuLeuProValSerArgTlleIleValHisProGlnPheTyrTlleIleGlnThrG1yAla 120
 DB 311 CTGCTGCGCGTCCACACAGGATCATGTGAGCCACACAGTTTCAATCATCCACAGCTGGAGCG 370

OY 121 AspIleAlaLeuLeuGlnLeuGlnGluProValAsnIleSerSerArgValHisThrVal 140
 DB 371 GATATCGCCCTGCTGAGCTGAGAGGCGCCGTGAACATCTCCAGCCGGGTCCACAGCGTTC 430

OY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTyr 160
 DB 431 AMGCTGCCCTGCTCGGAGACCTTCCCGGGGAGTGGCCGTGCTGGTCACTGGCGTGG 490

OY 161 G1yAspValAspAAspGluProleuProProPheProleuLysGlnVallyVal 180
 DB 491 GCGGATGTGAGCAATGATGAGCCCTCCAGCCGCAATTTCCCTGAAACAGATGAAAGTTC 550

OY 181 ProIleMeTcG1uAAsnHisIleCysAspAlaLysTyrHisIleuG1yAlaTyrThrG1yAsp 200
 DB 551 CCCATTAATGAAAACCACTTTGTGACGCAAAAATACCACTTGGCCCTCAKACGGGAGAC 610

OY 201 AspValArgIleIleArgAspAAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 611 GACGTCGGCATCATCCGTGACGACATGCTGTGTGCGGGGAAACAGCAGGAGGACTCATGC 670

OY 221 LysG1yAspSerG1yG1yProleuValCysLysValAsnG1yThrTrpLeuGlnAlaGly 240
 DB 671 CAGGGCAGACTCTGGAGGGCCCTGTGTGCAAGGTAATGCAACCTGGCTAACAGCGGGGC 730

OY 241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProG1yIleTyrThrArgVal 260
 DB 731 GTGGTCAGCTGGAGCGAGGGCTGTGCTCCAGCCCAACCGGCTGGCATCTACACCGGTTC 790

OY 261 ThrTyrTyrLeuAspTrpIleHisTyrValProLysLysPro 275
 DB 791 ACCTACTACTTGAGACTGATCCACCACTATGTCTCCCAAAAAGCCG 835

RESULT 4
 BC028059 1206 bp mRNA linear PRI 06-OCT-2003
 LOCUS Homo sapiens tryptase, alpha, mRNA (cDNA clone MGC:39869
 DEFINITION IMAGE:5181815), complete cds.

ACCESSION BC028059
 VERSION BC028059.1 GI:20380844
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 1206)
 Strausberg R.L., Fellingsold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Stenzel P., Prange C., Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H., Richards S., Worrley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Gimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalins D.B., Scherch A., Schein J.E., Jones S.J. and Marra M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
 PUBMED 22388257
 REFERENCE 2 (bases 1 to 1206)
 AUTHORS Strausberg R.
 TITLE Direct Submission
 JOURNAL TITLE Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc_mgc@nsl.nih.gov
 Akher N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.M., Bouffard G.G., Breen K., Brinkley C., Brooks S., Dietrich N.L., Granite S., Guan X., Gupta J., Haghghi P., Hansen N., Ho S.-L., Karlins E., Kong P., Larric P., Legaspi R., Maduro O.L., Masello C., Maskeri B., Mastrian S.D., McCloskey J.C., McDowell J., Pearson R., Stantrop S., Thomas S.J., Touchman J.W., Tsujieon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.

REMARK COMMENT
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRAC Plate: 62 Row: e Column: 8
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13699841.
 Location/Qualifiers
 1..1206
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:39869 IMAGE:5181815"
 /tissue="Brain, Image, Testis, adult, pooled whole"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

FEATURES source

gene
 1..1206
 /gene="TPS1"
 /note="synonym: TPS2, alpha II"
 /db_xref="LocustID:7176"
 /db_xref="MIM:191080"
 13..840
 /gene="TPS1"
 /codon_start=1
 /product="TPS1 protein"
 /protein_id="AAH28059.1"
 /db_xref="GI:20380845"
 /db_xref="locustid:7176"
 /db_xref="MIM:191080"
 /translation="MLSLILALPYLVAAPAAAPVQALQAGIVGGQEAPEARSKMPVQVSLRDRYMMHFCGSSLIHPQWVLTAAHCIGPVDKLAITRVQLREQHIYYQDQLPVSRIIVHPOFIIQADIALLELEEVNLSRVHTVMLPASPFPQMPQVITGWGDVNDRELPPEPPEPKVYVIMENHICDAAKYHGAALYVDDVRIIRDMILCAQNRDSCQDGGPFLVCKVNGTMDQAGVSVSDSCQAPNRPFGIYTYVYIYLDIHHYVPKP"
 103..813
 /gene="TPS1"
 /note="tryp. Spec. Region: Trypsin-like serine protease"
 /db_xref="CDD:smart0020"

misc_feature
 Alignment Scores:
 Pred. No.: 4,466-134 Length: 1206
 Score: 1494.00 Matches: 271
 Percent Similarity: 99.64% Conservative: 3
 Best Local Similarity: 98.55% Mismatches: 1
 Query Match: 98.81% Indels: 0
 DB: Gaps: 0

US-09-598-982c-52 (1-275) x BC028059 (1-1206)
 QY 1 MetLeuSerLeuLeuLeuLeuLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 DB 13 ATGCTGAGGCTTCCTGCTGCTGCGCTGCGCTGCGCTGCGCGACCGCGCTTACCGCGCCCT 72
 QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
 DB 73 GCCCAAGTCCAGGCCCTCGAGCAAGCGGGTATCGTGGGGGGTCAAGAGGCCCGCCAGGAGC 132
 QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 DB 133 AAGTGGCCCTGGCAAGTGAAGCTGAGAGTCCCGCCAGCATCTGATGCACCTTCGTGGGG 192
 QY 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaIleHisCysLeuGlyProAspVal 80
 DB 193 GGCCTCCCTCATCCACCCCGAGTGGGTGCTGACCCGCGCGCACTGCTGGACCGGACCGT 252
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisIleuTyrTyrGlnAspGln 100
 DB 255 AAGGATTCGGCCACCCCTCAAGGGTGCACATCGGGGAGCAGCACCTTACTACAGGACCAAG 312
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyVal 120
 DB 313 CTGCTGCGCGGTCAAGAGATCATCTGCAACCCACAGTTCATCATATCATCATGAGAGCG 372
 QY 121 AspIleAlaLeuLeuGlnLeuGlnGlnIleProValAsnIleSerSerArgValHisThrVal 140
 DB 373 GATATCGCCCTTCCTGAGCTGAGGAGGAGCCCGTGAACATCTCCACCGCGGTCCACACCG 432
 QY 141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrpValThrGlyTrp 160
 DB 433 ATGCTGCCCCCTGCTCGGAGACCTTCCCGGGGGAGTCCGCTGGGTCACTGCTGG 492
 QY 161 GlyAspValAspAsnAspGluProLeuProProPheProLeuIleGlnValIleVal 180
 DB 493 GCGGATGCGGCAAAAGATGAGGCCCTCCACCGCATTTCCCTGAAACAGGTGAAGGTG 552
 QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleuGlyValIleTyrThrGlyAsp 200
 DB 553 CCATTAATGGAACCAACCATTTTGTGACCGCAAAATATCAACTTGGCGCTTACAGGGAGAC 612

QY 141 MetLeuProProAlaSerGIUtrhrpProProGIWecProCysTrpValThrGIYTP 160
 DB 426 ATGCTGCCCCCTGCTGGGAGACCTTCCGCCGGGATGCGCTGGGCTCACTGGCTGG 485
 QY 161 G1yAspValAspAsnAspG1uProLeuProProProPherProLeuYsG1nValYlyVal 180
 DB 486 GGGGAGTGTGACATATATATGAGCCCTCCCAACCCTTTCCCTCTGAAGCAGGGTGAAGGTC 545
 QY 181 ProI1EmeG1uAsnHs11eCysAspAlaYsTYrH1sLeuG1yAlaTYrThrGIYAsp 200
 DB 546 CCGATATATGGAANAACAATTGTGTGAGGAAATACACTTGGCCCTTACACGGGAGAGC 605
 QY 201 AspValArg11e11eArgAspAspMetLeuCysAlG1yAsnSerG1nArgAspSerCys 220
 DB 606 GACGTCGGCATCCCGTGTGACATGCTGTGTGCGGGGAAACCCCGGAGGCTCATGTC 665
 QY 221 YsG1yAspSerG1yG1yProLeuValCysYsYsValAsnG1yThrTrIrrLeuG1nAlaG1y 240
 DB 666 CAGGGGACCTCTGAGAGGCCCCCTGGTGTGCAAGGTAAATGACACTGGCTACAGGGGGC 725
 QY 241 ValValSerTrpAspG1uG1yCysAlaG1nProAsnArgProG1y11eTYrThrArgVal 260
 DB 726 GTGGTACAGCTGGGACGAGGGCTGTGCCCCAGCCCAACCGGCTTGGCATTCACACCCGCTTC 785
 QY 261 ThrTYrTYrLeuAspTrp11eH1sH1sTYrValProLYsYsPro 275
 DB 786 ACCTACTACTTGGATCGATCCACCACTATATGTCCTCCCAAAAAGCCG 830
 RESULT 6
 BC074975 896 bp mRNA linear PRI 04-AUG-2004
 LOCUS BC074975 Homo sapiens tryptase beta 1, mRNA (cDNA clone MGC:103905
 DEFINITION IMAGE:30915294), complete cds.
 ACCESSION BC074975
 VERSION BC074975.2 GI:50960667
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1. Bases 1 to 896
 AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.R., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millan M.S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E., Scherch T.A., Schein J.E., Jones S.J. and Marra M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 896)
 AUTHORS Director MGC Project.
 TITLE Direct Submission
 JOURNAL Submitted (25-UN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT On Aug 4, 2004 this sequence version replaced gi:49901651.

Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Genome Sequence Centre, British Columbia Cancer Center
 CDNA Library Preparation: British Columbia Cancer Research Center
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Genome Sequence Centre
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Gunn, Nancy Liao, Kim Macdonald, Amara Maason, Mike R. Mayo, Josh Moran, Ryan Morin, Telka Olson, Diana Palimquist, Amca Petrescu, Anna Litsa Prabu, Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska, Duane Smallus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov
 Series: IRBU Plate: 2 Row: C Column: 7.
 Location/Qualifiers
 1..896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:103905 IMAGE:30915294"
 /tissue_type="Pooled tissue, PCR rescued clones"
 /clone_lib="NTR MGC 271"
 /note="Vector: pCR4_Topo TA with reversed insert"
 1..896
 /gene="TPSB1"
 /note="synonyms: TPS2, beta I, tryptase-II, TPS1"
 /db_xref="LocusID:7177"
 /db_xref="MIM:191081"
 34..861
 /gene="TPSB1"
 /codon_start=1
 /product="beta tryptase 1, precursor"
 /protein_id="AAH74975.1"
 /db_xref="GI:50960668"
 /db_xref="LocusID:7177"
 /db_xref="MIM:191081"
 /translation="MLSLLALPYLASRAYAARPGALORVGVGGQEPKSKMP
 QVSLRHHVPMWHPFGSSLIHPQWVTAHVRGPDVLDLALRVQLRQHLVYDQDL
 PVSRLIHPQFTYQIGDIALLELEBVNVSSVHTVLPASSTPPGMCWVTG
 GDVVDDELPPPLKQVPIVMEHHIDAKYHGAAYGDDVRYLRDMDLCAQVTRD
 SCQGGSPPLVCKVKGWTLQAGVYVSWGSGCAQPRRPGIYTVTYLIDWIIHHVPRKP"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.58e-126 Length: 896
 Score: 1414.00 Matches: 255
 Percent Similarity: 95.27% Conservative: 7
 Best Local Similarity: 92.73% Mismatches: 13
 Query Match: 93.52% Indels: 0
 DB: 9 Gaps: 0
 US-09-598-982c-52 (1-275) x BC074975 (1-896)
 QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTYrAlaAlaPro 20
 DB 34 ATGCTGAGCCCTGCTCTGCTGCGCTGCCCTCGGGAGCCGGCTTACGCGGCCCTT 93
 QY 21 AlaProValG1nAlaLeuG1nG1nAlaG1y11eValG1yG1nG1nAlaProArgSer 40
 DB 94 GCCCAAGGCCAGGCGCTTGCAGAGAGTGGCATCTGTGGGGGTCAAGAGCCGCCAAGAGC 153
 QY 41 YsTYrProTYrG1nValSerLeuArgValAlaArgAspArgTYrTrpMetH1sPheCysG1y 60
 DB 154 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCAGCGCCCACTGATGATCAGCTTCTGCGGG 213
 QY 61 G1ySerLeu11eH1sProG1nTrpValLeuThrAlaAlaH1sCysLeuG1yProAspVal 80

DB 214 GGCTCCCTCATTCCACCCCAAGGGTGTCTGACCCGAGGCGACTGGTGGACCCGACCTC 273
 QY 81 LysApPLeuAlaIaThrLeuArgValGlnLeuArgGlnIshLeuTyrTyrGlnAspGln 100
 DB 274 AAGGATCTGGCCCGCCCTCAGGGTGAACCTGGCCGGAGCAGCACTTACTACCCAGGACCCAG 333
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyVal 120
 DB 334 CTGCTGGCCGGTACGAGGATCTGTGACCCACAGTTCTAACCCGCCCAAGTCCGAGCG 393
 QY 121 AspIleAlaLeuLeuGlnLeuGlnIubProValIAsnIleSerSerArgValHisThrVal 140
 DB 394 GACATCCGCCCTCTGTGAGTGGAGGAGCCGGTGAACGTCCTCCAGCCAGTCCACACCGGTC 453
 QY 141 MetLeuProProAlaIleSerGluThrPheProProGlyMetProCysTyrValThrGlyTyr 160
 DB 454 ACCCTCCCGCCCTGCTCAGAGACCTTCCCGCCGGGAGATGCCGCTGGGTCACTGGCTGG 513
 QY 161 GlyAspValAspAsnAspGluIubProLeuProProPheProLeuLeuGlnValIubVal 180
 DB 514 GCGGATGGAGCAATGATGAGCGGCTCCACCCGCAATTTCTCTGAAGCAGGTGAAGGTC 573
 QY 181 ProIleMetGluAsnHisIleCysAspAlaIubTyrHisIleuGlyValIaTyrThrGlyAsp 200
 DB 574 CCCATTAATGGAAACACACATTTGTGACGGAAATACACACTTGGCCCTTACACCGGGAGAC 633
 QY 201 AspValAlaGleIleIaArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 DB 634 GACGTCGCCCATGTGCTGAGACGACATGCTGTGTGCGGGGAGAACCCCGGGAGGACATGTC 693
 QY 221 LysGlyAspSerGlyGlyProLeuValCysIubValAsnGlyTyrThrIubLeuGlnIaGly 240
 DB 694 CAGGGCGACTCCGGAGGGCCCTGGTGTGCAAGGTAAGGACCACTGGTGGAGCGGCGC 753
 QY 241 ValValSerTyrAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
 DB 754 GTGGTCACTGGGGGGAGGGGCTGTGCTCCAGCCCAACCGCCCTGATCTACACCCCTGTTC 813
 QY 261 ThrTyrTyrLeuAspTyrIleHisIshIshIshIshIshIshIshIshIshIshIshIsh 275
 DB 814 ACCCTACTTGTGACTGATCCACCACTATGTGTCCCAAAAAGCCG 858

RESULT 7
 BC074974
 LOCUS
 DEFINITION Homo sapiens tryptase beta 2, mRNA (CDNA clone MGC:104064
 IMAGE:30915512), complete cds.
 ACCESSION BC074974 GI:50959860
 VERSION BC074974.2 GI:50959860
 KEYWORDS MGC.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 896)
 Struhsberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Sherman,C.M., Schuler,G.D.,
 Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Martusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loughran,J.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richardson,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hilyk,S.W.,
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Panay,U., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,J.A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,A., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smallus,D.E.,

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT

Scherrch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 1247932
 2 (bases 1 to 896)
 Director MGC Project.
 Direct Submission
 Submitted (25-JUN-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 4, 2004 this sequence version replaced gi:49901868.
 Contact: MGC help desk
 Email: gcgbs-remail.nih.gov
 Tissue Procurement: Genome Sequence Centre, British Columbia Cancer
 Center
 cDNA Library Preparation: British Columbia Cancer Research Center
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palamquist, Anca Petrescu, Anna Liisa Prabh, U
 Parvaneh Saedi, JR Santos, Angeliq Scherrch, Ursula Skalska,
 Duane Smalflug, Jeff Stoltz, Miranda Tsai, George Yang, Jacquie
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNI at: <http://image.llnl.gov>
 Series: IRBU Plate: 4 Row: E Column: 9.
 Location/Qualifiers
 1..896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:104064 IMAGE:30915512"
 /tissue_type="Pooled tissue, PCR rescued clones"
 /clone_lib="NIH MGC 271"
 /note="Vector: pCR4_Topo TA with reversed insert"
 1..896
 /gene="TPSB2"
 /db_xref="LOCUSID:64499"
 34..861
 /gene="TPSB2"
 /codon_start=1
 /product="tryptase beta 2, precursor"
 /protein_id="AAH74974.1"
 /db_xref="GI:50959861"
 /db_xref="LOCUSID:64499"
 /translation="MLNLIILALPLVLSRAVAAPQALQRVGIVGGQEARSKWPV
 QVSRVHGVPMHFRGGSLIHPQWVLAARVGVGDVVDLAAIRVLRLEBOHLYYDOLL
 PVSIIYHPQYTRQIGADTALDELBRPVKYSVSHVHTVTLPAASETPRQMPQVITG
 GDVNDERLPPPLKQVFKVIMENHNLCDATYHLAGATYGDVRIVRDMLCAGMTRRD
 SCQDSDSGPIVCKVNWGLDAGVSVWBGCAQPNRPFGIYTRKTYLLDMIHRYVPKP"

gene
 CDS
 ORIGIN
 Alignment Scores:
 Fred. No.: 1.16e-125 Length: 896
 Score: 1405.00 Matches: 253
 Percent Similarity: 94.91% Conservative: 8
 Best Local Similarity: 92.00% Mismatches: 14
 Query Match: 92.92% Indels: 0
 DB: 9 Gaps: 0
 US-09-598-982C-52 (1-275) x BC074974 (1-896)
 QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaIaPro 20

Db 34 ATGCTGAATCTGCTGCTGGCGCTGCCCTGCTGCGAGACCGCCCTAACCGCGCCCT 93
 Qy 21 AAlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
 Db 94 GCCCCAGCCAGGCGCTGCGAGCGAGTGGGCGATGTTGGGGGTGAGGAGGCCCCAGGAGC 153
 Qy 41 IySTRPProTTPGlnValSerLeuArgValArgAspArgTyrTrpMetHispheCysGly 60
 Db 154 AAGTGGCCCTGGGAGGTGAGCTGAGCGATCCAGGAGTCCAGCCCACTGATGATGATGATG 213
 Qy 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisGlyIleuGlnProAspVal 80
 Db 214 GGCTCCCTCAATCCACCCCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 273
 Qy 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisSerLeuTyrTrpGlnAspGln 100
 Db 274 AAGGATCTGGCCCGCTCAGGAGTGCACATGCGGAGACAGCAGCAGCAGCAGCAGCAGC 333
 Qy 101 LeuLeuProValSerArgIleIleValHisSerGlnPheTyrIleGlnThrGlyAla 120
 Db 334 CTGCTCCCGGTCAAGAGATCATCTGCAACCCACAGTTCACACCCCGCCAGATCGGAGCG 393
 Qy 121 AspIleAlaLeuLeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 140
 Db 394 GACATGCGCTGCTGGAGCTGGAGAGCCGGTGAAGGATCTCCAGCCAGCAGCAGCAGCAG 453
 Qy 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTyrValThrGlyTyr 160
 Db 454 ACCCTCCCGCTGCTCAGAGACCTTCCCGCGGGAGATGCGTGGGTCACTGGCTGG 513
 Qy 161 GlyAspValAlaAspAspArgIleuProLeuProProPheProLeuLysGlnValLysVal 180
 Db 514 GGGGAGTGGAGCAATATGATGAGCCCTCCCAACCCCTTCCTGTGAAGAGGTGAAAGTTC 573
 Qy 181 ProIleMetGlnAspHisIleCysAspAlaLysTyrHisIleuGlnValAlaTyrThrGlyAsp 200
 Db 574 CCAATTAATGGAAACCAACATTTGTGACGGAAATACCACTTGGCCCTTACACGGGAGAC 633
 Qy 201 AspValArgIleIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
 Db 634 GACGTCGGCATGCTGCGTGCAGCAGCATGCTGTGGCGGGGAAACACCGGAGGGAGCTCAT 693
 Qy 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
 Db 694 CAGGGGAGCTCCCGAGAGGGCCCTGCTGTCAAGGTGAATGGCACCTGGCTGCGGGGCG 753
 Qy 241 ValValSerTrpAspGlnGlyCysAlaGlnProAspArgProGlyIleTyrThrArgVal 260
 Db 754 GTGGTACGCTGGGGCGAGGGGCTGTGCCCAAGCCCAAGGGCTGGCATTCACCCGCTGC 813
 Qy 261 ThrTyrTyrLeuAspTrpIleHisIleTyrValProLysLysPro 275
 Db 814 ACCTACTACTGTGAGTGCATCCCACTATGTCCTCCAAAAGGCCG 858
 RESULT 8
 HUMBT1RP HUMBT1RP 1143 bp mRNA linear PRI 31-OCT-1994
 LOCUS Human beta-tryptase mRNA, complete cds.
 DEFINITION M37488
 ACCESSION M37488.1 GI:179583
 VERSION beta-tryptase; mast cell serine protease; tryptase.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1143)
 MILLER, J.S., MOXLEY, G. and SCHWARTZ, L.B.
 Cloning and characterization of a second complementary DNA for
 human tryptase
 J. Clin. Invest. 86 (3), 864-870 (1990)
 JOURNAL MEDLINE 90369005
 PUBMED 2203827
 COMMENT Original source text: Human adult lung mast cell, cDNA to mRNA,

clone 10.
 Draft entry and computer-readable sequence for [J. Clin. Invest. (1990) In press] kindly submitted by L.B. Schwartz, 06-AUG-1990.
 Location/Qualifiers
 1..1143
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="16"
 /cell_type="mast cell"
 /issue_type="Lung"
 1..1143
 /gene="TPS2"
 /gene="TPS2"
 /product="beta-tryptase"
 6..833
 /gene="TPS2"
 /codon_start=1
 /product="beta-tryptase"
 /protein_id="AAA51843.1"
 /db_xref="GI:179584"
 /db_xref="GDB:600-125-891"
 /translation="MLNLLDLALPVLASRAYAARPAEQALQRVGIVGGQEARPRSKWPMQVSLRKHGPRYMMHFCGSSLHPQWVLTAAHCVGPDVLDLALRVQLRQHLHYDQQLPVSRIIVHPQETTAQIGADIALILELEEVKVSIVHTVTLRPSBTPPSPCWVTVGSDVNDERLPPFPKQVPLKIVMENHIDDAKYHIDAYGDVRLIVRDMLCAQVTRRDSQQGSGGRLVCKVNGTLDQAGVYVSWGSCAQPRRPGIYRVTYVLDIHHVPRKP"
 6..95
 /gene="TPS2"
 /note="G00-125-891"
 96..830
 /gene="TPS2"
 /product="beta-tryptase"
 /note="G00-125-891"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,54e-125 Length: 1143
 Score: 1405.00 Matches: 253
 Percent Similarity: 94.91% Conservative: 8
 Best Local Similarity: 92.00% Mismatches: 14
 Query Match: 92.92% Indels: 0
 DB: 9 Gaps: 0
 US-09-598-982c-52 (1-275) x HUMBT1RP (1-1143)
 Qy 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 Db 6 ATGCTGAATCTGCTGCTGGCGCTGCCCTGCTGCGAGACCGCCCTAACCGCGCCCT 65
 Qy 21 AAlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArgSer 40
 Db 94 GCCCCAGCCAGGCGCTGCGAGCGAGTGGGCGATGTTGGGGGTGAGGAGGCCCCAGGAGC 125
 Qy 41 IySTRPProTTPGlnValSerLeuArgValArgAspArgTyrTrpMetHispheCysGly 60
 Db 154 AAGTGGCCCTGGGAGGTGAGCTGAGCGATCCAGGAGTCCAGCCCACTGATGATGATGATG 185
 Qy 61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisGlyIleuGlnProAspVal 80
 Db 214 GGCTCCCTCAATCCACCCCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 245
 Qy 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisSerLeuTyrTrpGlnAspGln 100
 Db 274 AAGGATCTGGCCCGCTCAGGAGTGCACATGCGGAGACAGCAGCAGCAGCAGCAGCAGC 305
 Qy 101 LeuLeuProValSerArgIleIleValHisSerGlnPheTyrIleGlnThrGlyAla 120
 Db 334 CTGCTCCCGGTCAAGAGATCATCTGCAACCCACAGTTCACACCCCGCCAGATCGGAGCG 365
 Qy 121 AspIleAlaLeuLeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 140

DB	Accession	Score	Length	Alignment
Db	366 GACATCGCCCTGCTGAGCTGGAGGAGCCGGTGAAGGCTCTCCAGCCAGCTCCACACCGCTC	1405.00	1145	
Qy	141 MetleupProProAlaSerGluThrPheProProGlyMetProCysTTPValThrGlyTTP	94.91%	1145	
Db	426 ACCCTGCCCTGCTCAGAGACTTTCCCGGGGAGATGCGGCTGCTGCTGCTGCTGCTGCTG	92.00%	1145	
Qy	426 ACCCTGCCCTGCTCAGAGACTTTCCCGGGGAGATGCGGCTGCTGCTGCTGCTGCTGCTG	92.92%	1145	
Db	161 GlyAspValAspAsnAspGluProleuProProPheProleuGlyValIleValIleVal	92.92%	1145	
Qy	161 GlyAspValAspAsnAspGluProleuProProPheProleuGlyValIleValIleVal	92.92%	1145	
Db	486 GCGGATGAGCAAAATGATGAGCGCTCCACCGCATTTCTCTGAAACAGGTTGAAGGTTG	92.92%	1145	
Qy	181 ProIleMetGluAsnHisIleCysAspAlaIleValIleValIleValIleValIleValIle	92.92%	1145	
Db	546 CCCATTAATGAAAACCACTTTTGTATCCGAAATTCACCTTTGGCGCTTACAGCGGAGAAC	92.92%	1145	
Qy	201 AspValArgIleIleIleValAspAspMetLeuValIleValIleValIleValIleValIle	92.92%	1145	
Db	606 GACGTCGCCATGCTCGTGAAGCAATGCTGTGTCGCGGGAACACCCGGAGGGACTCATGTC	92.92%	1145	
Qy	221 LysGlyAspSerGlyGlyProleuValCysIleValIleValIleValIleValIleValIle	92.92%	1145	
Db	666 CAGGGCGACTCCGGAGGGCGCCCTGCTGTGCAAGTGAATGGCACCTGGCTGGAGCGGGGC	92.92%	1145	
Qy	241 ValValSerTTPAspGluGlyCysAlaIleGlnProAsnAspProGlyIleTTPThrArgVal	92.92%	1145	
Db	726 GTGTCACCTGGGGGAGGCGCTGTGCTCCAGCCCAACCGGCTGGAGTCTTACACCCGTGTC	92.92%	1145	
Qy	261 ThrTyrTyrLeuAspTTPIleHisIleTyrValIleValIleValIleValIleValIle	92.92%	1145	
Db	786 ACCACTACTTGGACTGGATCCACCACTATGTGCCCCAAAAGCCG 830	92.92%	1145	

DB	Accession	Score	Length	Alignment
Qy	555551	1405.00	1145	
LOCUS	beta-tryptase [human, basophil cell line KU812, mRNA, 1145 nt].			
DEFINITION	beta-tryptase [human, basophil cell line KU812, mRNA, 1145 nt].			
ACCESSION	555551			
VERSION	555551.1			
KEYWORDS	GI:265666			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens (human)			
REFERENCE	Blom, T. and Hellman, L. Characterization of a tryptase mRNA expressed in the human basophil cell line KU812. Scand. J. Immunol. 37 (2), 203-208 (1993)			
AUTHORS	Blom, T. and Hellman, L.			
TITLE	Characterization of a tryptase mRNA expressed in the human basophil cell line KU812			
JOURNAL	Scand. J. Immunol. 37 (2), 203-208 (1993)			
MEDLINE	93166209			
PUBMED	8434231			
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI g125219] from the original journal article.			
FEATURES	Location/Qualifiers			
source	1..1145			
gene	/organism="Homo sapiens"			
CDS	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	1..1145			
	/gene="beta-tryptase"			
	/note="precursor"			
	8..835			
	/gene="beta-tryptase"			
	/codon_start=1			
	/protein_id="AA013876.1"			
	/db_xref="GI:4261576"			
	/translation="MNLILLLALPVLASRAVAAPRQALQRFVGIIVGQEARSKMPVQVSLRHPYMMHPCGSLIHPQWVLTAAHCVKPVKDLAARVQLRQHLIYDQLPVSRIIVHPQFTQIGADILILESEVVKVSSHNTVTLRPASETFRPMLCPWVTGWGDVNDERLPPRPRLKQVTKVIRMEKHICDQKYLHAYTGDDVRIYRDMILCAQNTRRDSCQGSQGGELVCKVNGIWLQGVVSWGSCADQPNRPIYITRVYIILDIHTHIVPKP"			
ORIGIN				
Alignment Scores:	1.54e-125	Length:	1145	
Pred. No.:				

DB	Accession	Score	Length	Alignment
Qy	555551	1405.00	1145	
LOCUS	beta-tryptase [human, basophil cell line KU812, mRNA, 1145 nt].			
DEFINITION	beta-tryptase [human, basophil cell line KU812, mRNA, 1145 nt].			
ACCESSION	555551			
VERSION	555551.1			
KEYWORDS	GI:265666			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens (human)			
REFERENCE	Blom, T. and Hellman, L. Characterization of a tryptase mRNA expressed in the human basophil cell line KU812. Scand. J. Immunol. 37 (2), 203-208 (1993)			
AUTHORS	Blom, T. and Hellman, L.			
TITLE	Characterization of a tryptase mRNA expressed in the human basophil cell line KU812			
JOURNAL	Scand. J. Immunol. 37 (2), 203-208 (1993)			
MEDLINE	93166209			
PUBMED	8434231			
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI g125219] from the original journal article.			
FEATURES	Location/Qualifiers			
source	1..1145			
gene	/organism="Homo sapiens"			
CDS	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	1..1145			
	/gene="beta-tryptase"			
	/note="precursor"			
	8..835			
	/gene="beta-tryptase"			
	/codon_start=1			
	/protein_id="AA013876.1"			
	/db_xref="GI:4261576"			
	/translation="MNLILLLALPVLASRAVAAPRQALQRFVGIIVGQEARSKMPVQVSLRHPYMMHPCGSLIHPQWVLTAAHCVKPVKDLAARVQLRQHLIYDQLPVSRIIVHPQFTQIGADILILESEVVKVSSHNTVTLRPASETFRPMLCPWVTGWGDVNDERLPPRPRLKQVTKVIRMEKHICDQKYLHAYTGDDVRIYRDMILCAQNTRRDSCQGSQGGELVCKVNGIWLQGVVSWGSCADQPNRPIYITRVYIILDIHTHIVPKP"			
ORIGIN				
Alignment Scores:	1.54e-125	Length:	1145	
Pred. No.:				

ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 1194)

AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Magnet L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S., Lochellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Hale J.A., Gamaralle P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman D.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Matulis D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.

TITLE human and mouse cDNA sequences
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUBLISHED 2 (bases 1 to 1194)
REFERENCE 12477932
AUTHORS Strausberg, R.
TITLE Direct Submission

REMARK COMMENT
Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs_email.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLML)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sbgc.stanford.edu

Contact: (Dickson, Mark) med@paxli.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: http://image.lnl.gov
Series: IRAL Plate: 41 Row: J Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13775596.
Location/Qualifiers
1. .1194

FEATURES source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:32523 IMAGE:4695458"
/tissue_type="lung"
/clone_lib="NIH MGC_77"
/lab_host="DH10B"
/note="Vector: PDNR-LTB"
1. .1194

gene
/gene="TPS2B2"
/db_xref="LocusID:64499"
24_851
/gene="TPS2B2"
/codon_start=1
/product="Tyrosine beta 2, precursor"
/protein_id="AAH29356.1"
/db_xref="GI:20809955"
/db_xref="LocusID:64499"
/translation="MTLILILALPVLASRARAARPARQALQVIGYGGAPRSKRP
QVSLRVAHPTMWHFCGSGSLHPQWVILAAHCGPVDVADLALVQLRBOHLVYDQDL

CDS
/db_xref="LocusID:64499"
24_851
/gene="TPS2B2"
/codon_start=1
/product="Tyrosine beta 2, precursor"
/protein_id="AAH29356.1"
/db_xref="GI:20809955"
/db_xref="LocusID:64499"
/translation="MTLILILALPVLASRARAARPARQALQVIGYGGAPRSKRP
QVSLRVAHPTMWHFCGSGSLHPQWVILAAHCGPVDVADLALVQLRBOHLVYDQDL

ORIGIN
PVSRIIVHPQFTQAQIGADIALLEEREPVKSSHHHTVTLRPASETPRPMPWVW
GDVDDERLPPFPYKQVTPVIMENHIEDAKYTHIGDDVRIYRDMMLCAGNTRRD
SCQSGSGLPVCVKVGMWLDAGVSWGSGCAPNRPGLVYTVVYLDIMHIVPKRP"

Alignment Scores:
Prel. No.: 1 62e-125 Length: 1194
Score: 1405.00 Matches: 253
Percent Similarity: 94.91% Conservative: 8
Best Local Similarity: 92.00% Mismatches: 14
Query Match: 92.92% Indels: 0
Gaps: 0

US-09-598-982C-52 (1-275) x BC029356 (1-1194)

1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro
24 ATGCTGGAATCTCTGCTGCTGGCGCTGCCCGCTCCCTGGCGAGCGCGCTCAGCGCCCT
21 AlaProValAlaAlaLeuGlnGlnAlaAglY11eValIglYglYglnGlnAlaProArgSer
84 GCCCCAGGCCAGGCCCTTGAGCGAGTGGGAGTGGGGGTCAAGAGAGCCCGGAGAGC
41 LysTrrProTrrGlnValSerLeuArgValArgAspArgTrrTrrMetHisPheCysGly
144 AAGTGGCCCTGGCAAGTGAAGCTGAGAGTCCAGGCCCACTACTGATGACTTTCTGGCGG
61 GlySerLeuLeuHisPrrGlnTrrValLeuThrAlaAlaHisCysLeuGlyProAspVal
204 GGCCTCCATCATCACCCCAAGTGGTGTGACCGGAGGCACTGGTGGTGGAGCCGAGCGT
81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTrrGlnAspGln
264 AAGGATCTGGCCCGCTCGAGTGCAGCTGCGGAGGAGCCACTTACTACCAAGAGCAG
101 LeuLeuProValSerArgIleIleValHisPrrGlnPheTrrIleIleGlnThrGlyAla
324 CTGCTGCCGGTGAAGAGATCATCTGTGACCCCAAGTTCATACCGCCCAAGTGGAGCG
121 AspIleAlaLeuLeuGlnLeuGlnIleProValaenIleSerSerArgValHisThrVal
384 GACATCGCCCTGCTGAGTGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrrValThrGlyTrr
444 ACCCTGCCCTCCCTCGAGAGACTTCCCGGAGAGTCCCGGAGAGTCCCTGGGCTGAGCTGG
161 GlyAspValAspAsnAspGlnProLeuProProPheProLeuLeuGlnValLysVal
504 GCGGATGTGACAAATGATAGGCGCTCCACCGCAATTCCTGTAGAGCAAGGTGAAAGTTC
181 ProIleMetGlnAsnHisIleCysAspAlaLysTrrHisLeuGlyValaTyrThrGlyAsp
564 CCCATPAAAGAAACCAACATTTGTGACGGAAATPACACTTGGGGCCCTGACACCGGAGAGC
201 AspValAlaArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys
624 GACGTCGGCATGCTCGTGAAGCAATGCTGTGTGCGGGGAAACCCCGGAGGAGCAATGTC
221 LysGlyAspSerGlyGlyProLeuValCysLysValaenGlyThrTrrLeuGlnAlaGly
684 CAGGGGAGCTCGAGAGGCGCTGGTGTGCAAGGAAATGGCAACTGGGTGCAAGCGGGC
744 GTGGTCAAGTGGGGAGGAGGCTGTGTCCAGCCCAACCGGCTGGATCTACACCGCTGTTC
261 ThrTyrTrrLeuAspTrrPheHisIleTrrValProLysLysPro 275
804 ACCCTACTGACTGAGTCCAGTCCACCACTAATGTCCCAAAAAGCCG 848

RESULT 11
BC038416

Db 932 GCGGTCAAGTGGGGGAGGCGCTGTGCTCCAGCCCAACCGCGCTGCATCTACACCCGTGTTC 991
 QY 261 ThrTYrTYrLeuAspTrpIleHSHSHSYrValProlValSerPro 275
 Db ACCTACTACTGTGACTGATCCACCACTATGTCGCCCAAAAAGCCG 1036

RESULT 13
 CRS41836 825 bp mRNA linear PRI 29-JUN-2004
 LOCUS Homo sapiens full open reading frame cDNA clone RZPD0834B0932D for
 gene TPSB2, trypsase beta 2; complete cds, without stopcodon.
 ACCESSION CRS41836.1 GI:49456628
 VERSION Full ORF shuttle clone, Gateway(TM), complete cds.
 KEYWORDS Homo sapiens
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 825)
 AUTHORS Halleck, A., Ebert, L., Mkrandinya, M., Schick, M., Eisenstein, S.,
 Neubert, P., Ketrang, K., Schatten, R., Shen, B., Henze, S., Mar, W.,
 Korr, B., Zuo, D., Hu, Y., and Labaer, J.
 Cloning of human full open reading frames in Gateway(TM) system
 entry vector (pDONR201)
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 825)
 REFERENCE Halleck, A., Ebert, L., Mkrandinya, M., Schick, M., Eisenstein, S.,
 Neubert, P., Ketrang, K., Schatten, R., Shen, B., Henze, S., Mar, W.,
 Korr, B., Zuo, D., Hu, Y., and Labaer, J.
 AUTHORs Halleck, A., Ebert, L., Mkrandinya, M., Schick, M., Eisenstein, S.,
 Neubert, P., Ketrang, K., Schatten, R., Shen, B., Henze, S., Mar, W.,
 Korr, B., Zuo, D., Hu, Y., and Labaer, J.
 TITLE Direct Submission
 JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
 Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
 Germany
 COMMENT RZPD; RZPD0834B0932D, ORFN0 3770
 www.rzpd.de/cgi-bin/producte/cl.cgi?cloneID=RZPD0834B0932D RZPDLIB;
 Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
 834
 www.rzpd.de/cgi-bin/producte/showlib.pl.cgi?response=libno=834
 www.rzpd.de/products/orfclones/
 Contact: Inge Airlart
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel.: +49 30 32639 100
 Fax: +49 30 32639 111
 www.rzpd.de
 This clone is available from RZPD;
 Contact RZPD (customer.servic@rzpd.de) for further information.
 Clone name at Harvard Institute of Proteomics
 (www.hip.harvard.edu): RH130921.01L
 This CDS clone is part of a collection of human full ORF clones
 jointly established and verified by the Harvard Institute of
 Proteomics (HIP) and RZPD.
 This CDS has been cloned without stopcodon.
 The CDS has been inserted into pDONR201 via a BP Clonase(TM)
 reaction. Additional sequence has been added in front of the start
 codon: att. AAAAACA GCA GGC TCC ACC (ATG).
 The last codon is followed by the 3' att site: GACCCAGCTTTCTT. att
 The clone is validated by full sequence check.
 Compared to the reference sequence NM_003294 (GI:13775594) we found
 AA exchange(s) at position (first base of changed triplet):
 184 (ser->thr) 319 (ile->thr)
 Clone distribution: http://www.rzpd.de/products/orfclones/
 Location/Qualifiers

FEATURES
 source
 1. 825
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="RZPD0834B0932D"
 /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
 /lab_host="DH5A1pna"
 /note="Vector: pDONR201, Site_1: attPl, Site_2: attP2"
 1. 825
 gene

Alignment Scores:
 Pred. No.: 1.64e-125 Length: 825
 Score: 1403.00 Matches: 252
 Percent Similarity: 94.91% Conservative: 9
 Best Local Similarity: 91.64% Mismatches: 14
 Query Match: 92.79% Indels: 0
 Gaps: 0
 DB: 9
 US-09-598-982c-52 (1-275) x CRS41836 (1-825)

CDS
 /gene="TPSB2"
 1..>825
 /gene="TPSB2"
 /codon_start=1
 /protein_id="CA646635.1"
 /db_xref="GI:49456628"
 /translation="MLNLLDLLALPVLASRAVAAPACALORVGI VGGQEARPKMP
 QVSRVHGPKVMHRCGGSTLIRPMLVLAHCVGVDVMDLAIRVQLRQHLYYQDOL
 PVSRRTYHPPPTQITGADLALSLERPNVYSSVHTVTVLPASRTPRCPWPCVTVG
 GDVNDERLPPFPLKQVAFIMENHICDAKTYHIGATTGDDVRI VRDMDLCRGTBRD
 SCGGDSGPIVCKVMGTWLDAGVVSQWREGCAQPNRPGLIYTRVTVYLDIMIHRYVPRKQ"

QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTyAlaAlaPro 20
 Db 1 ATGCTGATTCGTCTGTCTCTCTGCGCGCTGCGCGCTCGTCCGGAGACCCTTACCGCCCTT
 QY 21 AlaProValAlaLeuAlaLeuGlnAlaAlaGlyTlleValGlyGlnGluAlaProArgSer 40
 Db 61 GCGCCAGGCGCCGCGCTCAGAGTGCGATCCGCGGCGGTCAAGGCGCCCGCGGAGCG
 QY 41 LysTrpProTrpGlnValSerLeuArgValAlaArgAspArgTrpTrpMetHisPheCysGly 60
 Db 121 AAGTGGCCCTGGAGTGAAGCTGAAGTCCACGGCCCACTGAGTACCTTCTGCGG 180
 QY 61 GlySerLeuIleHisProGlnTrpValIleuThrAlaAlaHisCysLeuGlyProAspVal 80
 Db 181 GGCACCCTCATCCACCCCAGATGGTGTGACCCGAGCGCACTGTGCGGACCGGACGCTC 240
 QY 81 LysAspLeuAlaThrLeuArgValAlaGlnLeuArgGlnHisIleuTYrTrpGlnAspGln 100
 Db 241 AAGGATCTGGCGCCCTCAGGTCGCACTGGGGGAGCGACCTCTACTACCGAGCCAGC 300
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTYrIleIleGlnThrGlyVala 120
 Db 301 CTGCTGCCGGTCAGGAGGACCATGCTGACCACCAAGTCTACACCCGCAATGCGAGCGG 360
 QY 121 AspIleAlaLeuLeuGlnIleGluIleProValAsnIleSerSerArgValHisThrVal 140
 Db 361 GACATCGCCCTGCTGAGAGTGGAGAGCGCGGTGAACCTCTCCAGCCAGCTCCACAGCGGTC 420
 QY 141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrpValIleThrGlyTrp 160
 Db 421 ACCCTGCCCCCTCTCAGAGACCTTCCCCGGGGGATGCCGCTGCGGTCACTGCGCGTGG 480
 QY 161 GlyAspValaAspAspArgGluProLeuProProPheProLeuValGlnValIleVala 180
 Db 481 GCGGCTGGGCAAMTAGAAGCGCTCCACCGCATTCTCTGAMGACAGTGAAGGTTC 540
 QY 181 ProIleMetGluAsnHisIleCysAspAlaIleTYrHisIleuGlyAlaTYrThrGlyAsp 200
 Db 541 CCCATTAAAGAAAACAATTTGTGACGCAAAAATACACACTTTGGCGCTTACAGGGAGAC 600
 QY 201 AspValaIleIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnAspAspSerCys 220
 Db 601 GACGTCGGCATCGTCGGTAGAGACATGCTGTGTGTGCGGGAACACCCGGAGGAGACTCAATGTC 660
 QY 221 LysGlyAspSerGlyGlyProLeuValCysIleValaAsnGlyTrpTrpLeuGlnAlaGly 240
 Db 661 CAGGGCGACTCCGAGGGCGCTGGTGGCAAGGTGAATGACACTGGCTGCAAGCGGGG 720
 QY 241 ValValaSerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIleTYrThrArgVal 260

Db 721 GTGGTAGCTGGGGCCAGGCGTGTGCCAGCCCAACCGGCTGGCATCTACACCCGTGTC 780

Qy 261 ThrTYrTYrLeuAspTRPLeuHisIstYrValProlyslvsPro 275
 |||
 781 ACCTACTACTGTGGACTGGANTCCACCACTATGTCTCCCAAAAAGCCG 825

RESULT 14

AR080460 AR080460 1137 bp DNA linear PAT 31-AUG-2000
 DEFINITION Sequence 13 from patent US 5968782.
 AR080460
 ACCESSION AR080460
 VERSION AR080460.1 GI:10007195
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 1137)
 AUTHORS Stevens, R.L.
 TITLE Mast cell protease that cleaves fibrinogen
 JOURNAL Patent: US 5968782-A 13 19-OCT-1999;
 FEATUERS Location/Qualifiers
 source 1.1137
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	2,98e-125	Length:	1137
Score:	1402.00	Matches:	252
Percent Similarity:	95.24%	Conservative:	8
Best Local Similarity:	92.31%	Mismatches:	13
Query Match:	92.72%	Indels:	0
DB:	6	Gaps:	0

US-09-598-982c-52 (1-275) x AR080460 (1-1137)

Qy 3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTYrAlaAlaProAlaPro 22
 ::::|
 3 AATCTGCTGCTGCTGGCGCTGCCCTCTGTGGCCAGCGCCCTTACCGCCCTGCCCCA 62

Db 23 ValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGlnAlaProArgSerIvsTRP 42
 |||
 63 GGGCAGAGCCCTGGAGAGAGTGGGAGATCGTGGGGGTGAGAGGCCCCCGAGAGCAAGTGG 122

Qy 43 ProTRGlnValSerLeuArgValArgAspArgTYrTRPMeHisIspHeCysGlyIysSer 62
 |||
 123 CCTGGCAGGGTGGAGTGGAGAGTCCAGGCCCACTATCGAGTGGACTTTCTGGGGGCTCC 182

Db 63 LeuIleHisProGlnTRPValLeuThrAlaAlaHisCysLeuGlyProAspValIvsAsp 82
 |||
 183 CTCATCCACCCCCAGTGGGTGGTGGACCCGACCTGCGTGGGACCCGACCTCAAGGAT 242

Qy 83 LeuAlaThrLeuArgValGlnLeuArgGlnGlnHisIleuTYrTYrGlnAspGlnLeuLeu 102
 |||
 243 CTGGCCGCTTCAAGGTGCACTGCGGAGAGCACTTACTTACAGAGCAAGTCTGCTGG 302

Db 103 ProValSerArgIleIleValHisProGlnIspHeTYrIleIleGlnThrGlyAlaAspIle 122
 |||
 303 CCGGTACAGCAGATCCGTCGTCGACCCAGTCTTACACCGCCCAAGATCCGAGAGGCAATC 362

Qy 123 AlaLeuLeuGlnLeuGlnGlyProValAsnIleSerSerArgValHisIsthValMetLeu 142
 |||
 363 GCCCTCTGGAGCTGGAGGAGCCGGTGAACGCTCCAGGCAACGCTCCAGCGGACCCCTG 422

Db 143 ProProAlaSerGlnThrPheProProGlyMetProCysETrpValIsthGlyTrpGlyAsp 162
 |||
 423 CCCCTGCTCAAGAGACTTCCCCCGGGGATCCGTTGGTCACTGGCTGGGGCGGAT 482

Qy 163 ValAspAsnAspGlyIvsProLeuProProProGlnPheProLeuIvsGlnValIvsValProIle 182
 |||
 483 GTGGACATGATGATGAGGGCCGTCACCGCCATTTTCCTGAAGCAGGTAAGGCCCCCAAA 542

Db 183 MetGlnAsnHisIleCysAspAlaIvsTYrHisLeuGlyAlaIvsTYrThrGlyAspAspVal 202

Db 543 ATGAAAAACCACTTTGGAGCAAAAATACCACTTGGCGCTTACACAGGAGACAGCGTC 602
 |||
 203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysIvsGly 222
 |||
 603 CCACATGCTCCGTGAGCAGCATGTGTGTGCGCGGAAACACCCGAGGAGACTCATGCCAAGGGC 662

Qy 223 AspSerGlyIvsProLeuValCysIvsValAsnGlyIsthTRPLeuGlnAlaGlyValIvsVal 242
 |||
 663 GACTCCGGAGGGCCCTGGTGTGGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTTC 722

Db 243 SerTRPAspGlnGlyCysAlaGlnProAsnArgProGlyIleTYrTYrThrArgValThrTYr 262
 |||
 723 AGCTGGGGCGAGGGTGTGGCCAGCCCAACCGCTGGCATCTACACCCGTGTCACCTAC 782

Qy 263 TYrLeuAspTRPLeuHisIstYrValProlyslvsPro 275
 |||
 783 TACTTGGACTGGATCCACCACTATGTCTCCCAAAAAGCCG 821

Db

RESULT 15

HUMTRRYLA 1137 bp mRNA linear PRI 03-AUG-1993
 LOCUS Human tryptase-I mRNA, 3' end.
 DEFINITION M33491
 ACCESSION M33491.1 GI:339980
 VERSION M33491.1 GI:339980
 KEYWORDS serine protease; trypase-I.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1137)
 Vanderslice, P., Ballinger, S.M., Tam, E.K., Goldstein, S.M.,
 Craik, C.S. and Caughey, G.H.
 Human mast cell trypase: multiple cDNAs and genes reveal a
 multigene serine protease family
 Proc. Natl. Acad. Sci. U.S.A. 87 (10), 3811-3815 (1990)
 90251647
 2187193
 Original source text: Human adult skin, cDNA to mRNA.
 Draft entry and computer-readable sequence for [1] kindly submitted
 by P.Vanderslice, 02-APR-1990.

COMMENT Location/Qualifiers
 1.1137
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="caxon:9606"
 <1..824
 /note="trypase-I"
 /codon_start=3
 /protein_id="AAA36778.1"
 /db_xref="GI:339981"
 /translation="MLILALPVLASRAYAAPADGQALQRVGIYGGDEAPRSKMPQV
 SLRVTYPMHRCGGSLILHPQVLTAAHCVDPVDLALALHQRBOHLIYQDILLPV
 SRIIVPQYTAOIGADIALLELSEPVVSHVHTVTLPPASEFPPEPMPGVGMD
 VDNDRLPPEPPIKQYKYPIMENHICDKAKHIGKATGSDVATVYDDMI CASNTPRDSC
 QGDSGRPLVCKVNGTWTLDAGVSVSWEBCAQENRREIYRVYIYVDMIHIVPKLP"
 1119..1124

ORIGIN polyA_signal

Alignment Scores:

Pred. No.:	2,98e-125	Length:	1137
Score:	1402.00	Matches:	252
Percent Similarity:	95.24%	Conservative:	8
Best Local Similarity:	92.31%	Mismatches:	13
Query Match:	92.72%	Indels:	0
DB:	9	Gaps:	0

US-09-598-982c-52 (1-275) x HUMTRRYLA (1-1137)

Qy 3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTYrAlaAlaProAlaPro 22
 ::::|
 3 AATCTGCTGCTGCTGGCGCTGCCCTCTGTGGCCAGCGCCCTTACCGCCCTGCCCCA 62

QY 23 ValGlnAlaLeuGlnGlnAlaGlyTyrLeuValGlyGlnGlnAlaProArgSerLysTrp 42
 DB 63 GGCCAGGCCCTGCAAGGTGGGCAATCGTGGGGGTCAAGAGGCCCCCAAGCAAGTGG 122
 QY 43 ProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySer 62
 DB 123 CCTGGCAGGTGATGCTGAGATCCACCGGCCCATATCTGATGCACTTCTGCGGGGGCTCC 182
 QY 63 LeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82
 DB 183 CTCATCCACCCCAAGTGGGTGTGACCCGACGGCACTGCGTGGGACCCGGACCTCAAGGAT 242
 QY 83 LeuAlaThrLeuArgValGlnLeuArgGlyGlnHisIleuTyrTrpGlnAspGlnLeuLeu 102
 DB 243 CTGGCCGCCCTGAGGGTCACTGGGAGACACTTACTACTACCGAGCAAGCTGTG 302
 QY 103 ProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyValAspIle 122
 DB 303 CCGGTACAGCAGATCATCTGCAACCCACAGTTCATACCCGCCCAAGATCGGAGCGGACATC 362
 QY 123 AlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisThrValMetLeu 142
 DB 363 GCCCTGCTGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACAGTCCACCGGTCACCCCTG 422
 QY 143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
 DB 423 CCCCCTGCTCAGAGACTTCCCCCGGGAGTCCGCTGGGATCACTGGGGGCGAT 482
 QY 163 ValAspAsnAspGluProLeuProProPheProLeuLysGlnValLysValProIle 182
 DB 483 GTGGACAAATGATGAGGGCTCCCAACCGCCATTTCTCTGAAAGCAAGTGAAGSTCCCCATA 542
 QY 183 MetGluAsnHisIleCysAspAlaIysTyrHisIleuGlyAlaTyrThrGlyAspAspVal 202
 DB 543 ATGGAAAACCACTTTGTGACCGCAAAATACCAACTTGGCGCCTTACACGGGAGACGACGTC 602
 QY 203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGly 222
 DB 603 CGCATCGTCCTGACGACATGCTGTGTGCCGGGAAACACCCGGAGGACTCATGCCAGGGGC 662
 QY 223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
 DB 663 GACTCCGGAGGGGCCCTGCTGTGCAAGTGGCACTGGCTGCAAGGGCCGGCTGTGTC 722
 QY 243 SerTrpAspGlnGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
 DB 723 AGCTGGGGGAGGGCTGTGCTCCCAAGCCAGCCGCGCTGGCATCTACACCCGCTCACCTTAC 782
 QY 263 TyrLeuAspTrpIleHisIleTyrValProLysLysPro 275
 DB 783 TACTTGGACTGATCCACCACTATGTCCCAAAAAGCCG 821

Search completed: August 30, 2005, 18:42:00
 Job time : 3564 secs

Third Page Blank (usps)

Gencore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 15:00:06 ; Search time 3116 Seconds
(without alignments) 3359.331 Million cell updates/sec

Title: US-09-598-982c-52

Perfect score: 1512

Sequence: 1 MSLIILLALPVLASRAVAAP.....ITRYTYVLDIMHHVVKPK 275

Scoring table:
BLOSUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delpop 6.0, Delpext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09598982/crunat_29082005_080946_26105/app_query.fasta_1.455
-DB=EST -OPMT=fasta -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=humand0.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCKL
-OUTFMT=ptc -NOM=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982@cgn2_1.5180@crunat_29082005_080946_26105 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEBUFFER -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est1:*
9: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match Length, DB ID, Description. Contains 9 rows of search results.

Table with columns: LOCUS, DEFINITION, ACCESSION, VERSION, KEYWORDS, ORGANISM, RESULT 1, ALIGNMENTS. Contains detailed information for the top hit (BM919268).

FEATURES source
1. 1030
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5748499"

```

/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/node="Organ: pooled pancreas and spleen; Vector:
PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

```

```

ORIGIN
Alignment Scores:
Pred. No.: 7,64e-133 Length: 1030
Score: 1463.00 Matches: 270
Percent Similarity: 98.92% Conservatve: 4
Best Local Similarity: 97.47% Mismatches: 1
Query Match: 96.76% Indels: 2
DB: 5 Gaps: 0
US-09-598-982C-52 (1-275) x BW919268 (1-1030)

```

```

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaPro 20
DB 33 ATGCTAGAGCCCTGCTGCTGGGGCTCCCTGCTGGGAGCCCGCCCTTACCGGGCCCT 92
QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyLeValGlyGlnGlnAlaProArgSer 40
DB 93 GCCCCAGTCCAGCCCTCGAGCAAGGGGGTATGTTGGGGGTCAAGAGGCCCCCAAGAAC 152
QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
DB 153 AAGTGGCCCTGGGAGGTGAGCTGAGAGTCCCGGAGCCGACCGATCGATGATGACCTTCTGGG 212
QY 61 GlySerLeuLeuHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
DB 213 GGCTCCCTATCCACCCCAAGTGGGTGCTGACCCGGCCGACCTCGGAGCCGGACGTC 272
QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTrpGlnAspGln 100
DB 273 AAGGATCTGGCCACCTCTCAGGGTGGCACTCGGGAGCAAGCACTTACTACAGACCCAG 332
QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
DB 333 CTGCTCCCGGTGAGCATCATCGTCAACCCCAAGTTCATCATCCAGACATGGAGCG 392
QY 121 AspIleAlaLeuLeuGlnLeuGlnGlnProValAsnIleSerSerArgValHisTrpVal 140
DB 393 GATATGCCCTGTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCTCCACACGGTTC 452
QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
DB 453 ATGCTGCCCTCCCTCGGAGACCTTCCCTCCCGGGGATCCGCTGGGTCACTGGCTGG 512
QY 161 G1YAspValAspAsnAspGlnProLeuProProProPheProLeuLysGlnValIlyVal 180
DB 513 GCGGATGTGAGCAATGATGAGCCCTCCACCGCCATTTCCCTGAAGGACGAGGAAAGTTC 572
QY 181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisIleGluValaLayrThrGlyAsp 200
DB 573 CCAATATGGAACCAACATTTGTGACGCAAAATACCACTTGGCCCTTACACGGGAGAC 632
QY 201 AspValArgIleLeuArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
DB 633 GACGTCCGCAATCTCCGTGACAGCATGCTGTGTGCGGGGAAACACCCGGAGGGACTATCC 692
QY 221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
DB 693 CAGGGGAGACTCTGGAGGGCCCTGCTGTGGAAGGTAAATGGGACCTGGGTCAACAGGGGGC 752
QY 241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrTrpArgVal 260

```

```

DB 753 GTGCTCAGCTGGGAGCAGAGGGGCTGTGCCCAAGCCCAACCCGCTGGCATCTACCCCTGT 812
QY 260 a1ThrTyrTrpLeuAspTrpIleHisIleTyrValProLysIlyAspPro 275
DB 813 TACCTACTACTTGGACTGATGATCCACCACTATGTCCCAAAAAGCCG 859

```

```

RESULT 2
BQ720404
LOCUS
DEFINITION BQ720404.1 bp mRNA linear EST 16-JUL-2002
IMAGE:6196776.5', mRNA sequence.
ACCESSION BQ720404.1 GI:21859301
VERSION BQ720404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-rmail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: LIML13605 row: m column: 01
High quality sequence stop: 587.
Location/Qualifiers

```

```

FEATURES
source
1..904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:6196776"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sympathetic_trunk"
/node="vector: PCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
directionality cloned using the following adaptors:
5'-TCGACCAAGCTCCG-3' and
5'-GACTAGTCTAGATCGAGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.5e-107 Length: 904
Score: 1204.50 Matches: 225
Percent Similarity: 88.06% Conservatve: 11
Best Local Similarity: 83.96% Mismatches: 21
Query Match: 79.66% Indels: 11
DB: 5 Gaps: 1
US-09-598-982C-52 (1-275) x BQ720404 (1-904)

```

```

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaPro 20
DB 10 ATGCTAGATCTGCTGCTGGGGCTCCCTGCTGGGAGCCCGCCCTTACCGGGCCCT 69
QY 21 AlaProValGlnAlaLeuGlnGlnAlaGlyLeValGlyGlnGlnAlaProArgSer 40
DB 70 GCCCAGGCCAGGCCCTCGAGAGTGTGTTGGGGGTCAAGAGGCCCCCAAGAGC 129

```

```

41  LysTrpProTrrGlnValSerLeuArgValArgAspArgTrpMetHisPheCysGly 60
130  NAGTGGCCCTGGCAAGTGGCTGTAAGTCCCGCCGACCCGATCTGATGATGATGATCTTCTGGCGGG 189
61  GlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
190  GAGCTCCCTCATTCCACCCCGAGTGGGTGCGACCGGACGGGACCTGCGTGGGAGACCG----- 243
81  LysAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
244  -----GTGCAACTGCCCGGGGAGCGACACTTACTACCAGGACCG 282
101  LeuLeuProValSerArgIleIleValHisProGlnPheTyrlIleGlnThrGlyVala 120
283  CTGCTCCCGGTGACGAGATCATCTGTGACCCACGATTTCTTACACCAGCCCAAGATCGGAGCG 342
121  AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
343  GACATCCCGCTGCTGAGGCTGGAGGCGCGGTGAACTGCTTCCAGCGACAGTCCGACACCGGTC 402
141  MetLeuProProValSerGluThrPheProProGlyMetProCysTrpValThrGlyTyr 160
403  ACCCTGCCCCCTCCAGACCTTCCCTCCCGGGGATGCCGTGCTGGGTGCTGCTGG 462
161  GlyAspValAspAspAspGluProLeuProProProPheProLeuLeuGlnValIlysVal 180
463  GCGGATGCTGGACAAATGATAGCGCCCTCCACCCGCAATTTCTTAAAGGAGGTGAAGGTC 522
181  ProlIleMetGluAsnHisIleCysAspAlaValTyrlHisLeuGlyValaTyrThrGlyLys 200
523  CCCAATATGGAATAACCAACTTGTGTGACGGAAAATTCACCTTGTGGCGCCCTACAGCGGGAAG 582
201  AspValAlaArgIleIleIleArgAspAspMetLeuCysAlaGlnProAsnArgProGlyIle 220
583  GACGTCCTCGGATCCCGTGAAGCAATGCTGTGGTCGGCGGGAACACCCGGGAGGACATCATCC 642
221  LysGlyAspSerGlyGlyProLeuValCysIlysValAsnGlyThrThrLeuGlnIleIleGly 240
643  CAGGCGGACTCCGAGAGGCGCCCTGGTGGGGAACGGGAAATGGCCCTGGGTGGCAGCAGGGGG 702
241  ValIValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThr-ArgV 260
703  GTGGGTCACTCTGGGGCGAGGGGCTGGGGCCCGGCCCAACCGGGGGGGGCTTTAACCCTGGGG 762
61  aLThrTyrTyrLeuAspTyr 266
763  CCCCCTCTACTGGGAATGG 782

```

```

RESULT 3
CD671851 619 bp mRNA linear EST 24-JUN-2003
LOCUS Fg07g11.y1 Human Iris cDNA (Normalized): Fg Homo sapiens cDNA clone
DEFINITION Fg07g11.y1, mRNA sequence.
ACCESSION CD671851
VERSION CD671851.1 GI:32173582
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 619)
Mistow,G., Bernslein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
Touhman,J.W., Bouffird,G., Smith,D. and Peterson,K.
Expressed sequence tag analysis of adult human iris for the NIBank
Project: steroid-response factors and similarities with retinal
pigment epithelium
Mol. Vis. 8 (4), 185-195 (2002)
JOURNAL MEDLINE PUBMED
22103462 12107412
Contact: Mistow G
Section on Molecular Structure and Function
National Eye Institute

```

```

6/3/31, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 07 row: 9 column: 11
Seq primer: M13R1 reverse primer (ABI) .
Location/Qualifiers
1..619
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Fg07g11"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMD108"
/clone_lib="Human Iris cDNA (Normalized): fg"
/note="Organ: Eye; Vector: pCMVSPORT6; A human iris library (hx) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (50C 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."
```

```

ALIGNMENT Scores:
Origins:
Pred. No.: 2.27e-100 Length: 619
Score: 1129.00 Matches: 202
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 98.54% Mismatches: 0
Query Match: 74.67% Indels: 0
DB: Gaps: 0
US-09-598-982C-52 (1-275) x CD671851 (1-619)

```

```

41  GlyGlnGluAlaProAspSerIleTrpProTrrGlnValSerLeuArgValArgAspArg 53
34  GGTAGAGAGGCCCCCAGGAGCAGAGTGGGCGGAGGTGAGAGTGCCGGGACCGCA 62
3  GGTAGAGAGGCCCCCAGGAGCAGAGTGGGCGGAGGTGAGAGTGCCGGGACCGCA 62
54  TrrTrpMetHisPheCysGlyGlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 73
63  TACTGATGACTCTTCTGTGGGGGCTCCCTCATCCACCCCGAGTGGGTGATGACCGCGCG 122
74  HisCysLeuGlyProAspValIlysAspLeuAlaThrLeuArgValGlnLeuArgGlnGln 93
123  CACTGCCCGGAGCCGAGGTCAGAGATCGGCACTCCGAGGGGACACCTGGGAGGAG 182
94  HisLeuTyrTyrlAspGlnLeuLeuProValSerArgIleIleValHisProGlnIle 113
183  CACTCTTACTTCCAGAGACTGCTGCTGCGGTGACAGATCATCTGCACACCAGATTC 242
114  TyrlIleIleGlnThrGlyValAlaAspIleAlaLeuLeuGluLeuGluGluProValAsnIle 133
243  TACATCATCCAGACTGAGGCGGATATAGCCCTGCTGGAGCTGGAGAGCCCGTGAACAATC 302
134  SerSerArgValHisIsthValMetLeuProProAlaSerGluThrPheProProGlyMet 153
303  TCCAGCCCGCGGCAACCGGTATCTGCTCCCTCCGAGGACCTTCCCGGGGAGTGG 362
154  ProCysTrpValThrGlyTrrGlyAspValAspAspAspGluProLeuProProPhe 173
363  CCGTCTGGGTCACTGGGTGGGGGAGTCGACATATGAGCCCTTCCACCGGCANTTT 422
174  ProLeuTyrGlnValIlysValProIleMetGluAsnHisIleCysAspAlaValTyrHis 193
423  CCCCAGAGCAGGTGAGAGTCCCATATATGAGAAAACCAATTGTGACCGCAAAAATACCA 482

```

QY 194 LeuGIYAAlTYrThrGIYAAspAaPValaIrgIleIIeAArgAspAaPMeIeuCySaIagIy 213
Db 483 CTGGCCGCTACAGGGGAGAGAGCACTCCGCATCATCCGGACAGCACTGCTGTGCGCGG 542
QY 214 AmsSerGIhArYAaSPSerCYsIySGIYAaSPSerGIyGIyProIeuValCYsIySAVaIaAn 233
Db 543 AACACCCGGAGGACATCATGCAGCGAGCCTGTGAGAGGGCCCTGTGTGCAAAAGTGAAAT 602
QY 234 G1YThrTriPleuGIa 238
Db 603 GGCACCTGGCTACAG 617

RESULT 4
CV030374 631 bp mRNA linear EST 20-AUG-2004
LOCUS 9476 Full Length cDNA from the Mammalian Gene Collection Homo
DEFINITION sapIens cDNA 5 similar to BC028059, mRNA sequence.

ACCESSION CV030374
VERSION CV030374.1 GI:51488624
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
AUTHORS Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Pauch,C.,
Vandenhaute,J., Cusick,M.E., Albaladejo,S., Hill,D.E. and Vidal,M.
Human ORFome Version 1.1: a Platform for Reverse Proteomics.
Genome Res. (2004) In press
CONTACT: Vidal M

COMMENT Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu

ORF Sequence Tag (OSTP) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGCTGAGCGCTGCTGCTG
BACKWARD: CACGGCTTGGGGACADVA
Insert Length: 631 Std Error: 27.00
Plate: 11081 row: 06 column: C
Seg primer: ACTGCGCTGCTGTTTACACACTGTCGTGACTGGGAAAC
High quality sequence start: 98
High quality sequence stop: 630
POLYA-No.

FEATURES
source
1..631 Location/Qualifiers
/origins="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/feature_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/note="Vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference: MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-Length Human and Mouse CDNA Sequences. PNAS,
2002, 99(26), 16899-16903"

Alignment Scores:
Prod. No.: 2 24e-99 Length: 631
Score: 119.00 Matches: 208
Percent Similarity: 99.05% Conservative: 0

Best Local Similarity: 99.05% Mismatches: 2
Query Match: 74.01% Gaps: 0
DB: 7 Indels: 0
Gaps: 0

US-09-598-982c-52 (1-275) X CV030374 (1-631)
QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTYrAla1a1aPro 20
Db 1 ATGTGAGGCTGCTCTCTCTCTCTGCGCTGCGCCGCTGGAGGCGCCGGCTACCGCCCTT 60
QY 21 A1aProValGlhAlaLeuGIhAlaIyIleValaIyGIyGIyGInG1a1a1aProArgSer 40
Db 61 GCCCAGTCCAGGCTTGCAGAAAGGGGGATTCGTGGGAGGTCAAGAGAGCCTCCAGAGCC 120
QY 41 IySTrPPrOTrGInValSerLeuArgValArGaPaArGTYrTTrMetHisPheCYsGIy 60
Db 121 AAGTGCCTTGGCAGGTAGAGCCTGTGAGATCCCGACCGATATCTGGATCACATCTGTGGG 180
QY 61 GIYSerLeuIIeHisPProGIhITrPValLeuThraAlaIahisCYaLeuGIYProkApVal 80
Db 181 GGCTCCCTCATCCACCCACAGGTGGGCTGAACCGCGCGCACTGCTGGAGACCGGACCGTCC 240
QY 81 IySApPLeuAlaThrLeuArgValGlhIeLeuArgGIyGInhiSIeUyTYrTYrGlnhAaSpGIh 100
Db 241 AAGGATCTGGCCACCTCAGGGGTGCACTGGAGAGAGGCAKACTTACTACTACAGGAGAG 300
QY 101 LeuLeuProValSerArgIleIleValhiSPProGIhPheTYrIleIleGIhThrGIyAla 120
Db 301 CTGGCTGGCGGTAGAGGATATCATGTGACCACCAAGTTCATCATCATCACAGCTGGAGCG 360
QY 121 AaPlIeAlaLeuLeuGIhLeuGIhUeGIhUeProVala1a1eSerSerArgValhiPThrVal 140
Db 361 GAATATCGCCCTGCTGAAGCTGGAGAGGCCGTGAACATCTCAGCGCGGCTCACACCGGTCC 420
QY 141 MetLeuProProAlaSerGIhThrPheProProGIYmetProCySTrTTrValThrGIyThrP 160
Db 421 ATGCTGCCCTCGCTGAGAGCTTCCCCCGGGAGAGCGGTGGGTGATCAGCTGGCTGG 480
QY 161 GIYAaPValAspAsnAspGluProLeuProProPheProLeuIyGlnValIyLeuVal 180
Db 481 GGCGATGTGGACAAGATGAGAGCCTCCACCGCATTTCCCTGAAGAGAGAGTGAAGGTC 540
QY 181 ProIleMetGluAenhiSIleCyAspAlaIySTYrHisLeuGIyAlaTYrThrGIYAsp 200
Db 541 CCATTAATAGAAAACACATTTGTGAGCCANNAATACCACTTTGGCGCTACAGCGGAGAC 600
QY 201 AsPValArgIleIleArgAspAaPMeIeu 210
Db 601 GACGTCCGCATCATCCGTGAGAGCAATGCTG 630

RESULT 5
CD616760 673 bp mRNA linear EST 12-JAN-2004
LOCUS 56015240J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD616760
ACCESSION CD616760.1 GI:40265024
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITILE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT CONTACT: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source Location/Qualifiers
1..673

ORIGIN

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"

```

Alignment Scores:

Pred. No.:	4,31e-99	Length:	673
Score:	1116.50	Matches:	216
Percent Similarity:	96.00%	Conservative:	0
Best Local Similarity:	96.00%	Mismatches:	8
Query Match:	73.84%	Indels:	4
DB:	6	Gaps:	0

US-09-598-982c-52 (1-275) x CD616760 (1-673)

```

OY 12 LeuAlaSerArgAlaTyrAlaAlaProAlaProValAlaLeuGlnAlaGlyIle 31
DB 2 CTGGCGAGCCCGGCTTACGGCCCTGCTCCAGTCCAGGCTCGAGAAAGCGGATTC 61
OY 32 ValGlyGlyGlnGlnAlaProArgSerIlyTTPProTTPGlnValSerLeuArgValArg 51
DB 62 GTGGGGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGC 121
OY 52 AAPPATGTYTTPMetHisPheCysGlyGlySerLeuIleHisProGlnTTPValLeuThr 71
DB 122 GACCCGATACCTGAGATGACTTCTGGGGGGCTCCCTCATCACCCCAAGTGGTGTGAC 181
OY 72 AlaAlaHisCysLeuGlyProAspValIlyAspLeuAlaThrLeuArgValGlnLeuArg 91
DB 182 GGGGGCACTGCTGGGAGCCGAGCCTGCAAGATCTGGCCACCCTCGAGGTCGAACCTGGCG 241
OY 92 GlnGlnHisLeuTyrTyrGlnHisProGlnLeuLeuProValSerArgIleIleValHisPro 111
DB 242 GAGCAGCACTTACTACTACAGAGACCGTGTGCTGCGCAGCAGAGATCATCGTGCACCCA 301
OY 112 GlnPheTyrIleIleGlnThrGlyValAspIleAlaIleLeuGlnLeuGlnIleProVal 131
DB 302 CAGTTCCTACATCATCCAGACTGAGGCGATATCGCTCTGAGCTGAGGAGCCCGCG 361
OY 132 AsnIleSerArgValHisThrValIleMetLeuProProAlaSerGlnIleThrPheProPro 151
DB 362 AACATCTCCAGCCGGCTCCACAGGTCATGCTGCCCTCGCGAGACCTTCCCGCG 421
OY 152 GlyMetProCysTTPValIleThrGlyTTP-GlyAspValAspAspAspIleProLeuProPr 171
DB 422 GGGATGCGGTGGGTGCTGCTGGGGGCGATGTGGCAAAATGATGAGACCCCTCCACAC 481
OY 171 GProPheProLeuIlyGlnValIlyValIleProIleMetGluAsnHisIleCysAspAlaIly 191
DB 482 GCMAATTTCCCGAAGCAGAGTGAAGGTCCCATTAAGAAAACAATTTGTGAGCGCAAA 541
OY 191 GTYRHisLeuGlyAlaIleTyrThrGlyAspAspValIleArgIleIleIleAspAspMetLeuCy 211
DB 542 ATRCCACTTTGGCGCTTACCGGAGACAGACTCCCGCATATCCGTGACGACATGCTGTG 601
OY 211 AsIaGlyAsnSerGlnArgAspSerCysIlySerGlyAspSerGlyGlyProLeuValCysIly 231
DB 602 TCGCCGGGAAACAG-CAGAGGAGCATC-TGCAGAGGGCGACTGTGGAGGGCCCTGTGTGACG 659
OY 231 sValAsnGlyThr 235
DB 660 -GTGAATGGCACC 671

```

RESULT 6 CD616759/c 682 bp mRNA linear EST 12-JAN-2004
LOCUS D616759.1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD616759
ACCESSION CD616759.1 GI:40265023
VERSION EST.
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (Bases 1 to 682)
AUTHORS Fu,G.R., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES Location/Qualifiers
source 1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/notes="Vector: pDrive Cloning Vector"

ORIGIN

```

Alignment Scores:
Pred. No.: 2.56e-97 Length: 682
Score: 1098.50 Matches: 221
Percent Similarity: 97.37% Conservative: 1
Best Local Similarity: 96.93% Mismatches: 5
Query Match: 72.65% Indels: 6
DB: 6 Gaps: 0

US-09-598-982c-52 (1-275) x CD616759 (1-682)
OY 9 LeuProValIleuAlaSerArgAlaTyrAlaAlaProAlaProValAlaLeuGlnGln 28
DB 678 CTGCCCCGTCCTGGGGGAC-C-CGGCCCTACCGGGCCCTGCCCCAGTCCAGGCTCGAGCA 620
OY 29 AlaGlyIleValGlyGlnGlnAlaProArgSerIlyTTPProTTPGlnValSerLeu 48
DB 619 GCGGGTATCTCGGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGGCAGAGTGCCTG 560
OY 49 ArgValAspAspArgTyrTTPMet-HisPheCysGlyGlySerLeuIleHisProGlnTTP 68
DB 559 AAGATCCCGCAACCACTATGATGACTTCTGGGGGGGCTCCCTCATCCACCCAG-TG 501
OY 68 PValLeuThrAlaAlaHisCysLeuGlyProAspValIlyAspLeuAlaThrLeuArgVal 88
DB 500 GGTGCTGACCCCGGGGCACTGCTGGGACCCGAGCCTCAAGATATGCGCACCTCAAGG 441
OY 88 GlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleI 108
DB 440 GCMACTGGGAGAGCACTTACTACTACAGAGACAGACTGCTGCCAGTCCAGCAGAGATCAT 381
OY 108 eValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGlnLeuGly 128
DB 380 CGTGCACCCACAGTTCATCATCATCAGACTGAGAGCGGATATGCGCCCTGTGAGCTGGA 321
OY 128 uGlnProValAsnIleSerArgValHisThrValIleMetLeuProProAlaSerGlnTh 148
DB 320 GAGAGCCCGTGAACATCTCCAGCCCGGTCACACAGGTCATGCTGCCCTCGTGGAGAC 261
OY 320 rPheProProGlyMetProCysTTPValIleThrGlyTTPGlyAspValAspAspAspIlePr 168
DB 260 CTTCCCCCGGGGAGATGCGGTGTGGGTCACTGGCTGGGGCCAGATGTGGACAAATGATGAGCC 201
OY 168 OLeuProProProPheProLeuIlyGlnValIlyValIleProIleMetGluAsnHisIleCy 188
DB 200 CTTCCCAACCCCAATTTCCCGAAGCAGAGTGAAGGTCCCATTAATGAAAACAATTTG 141
OY 188 sAspAlaIlysTyrHisLeuGlyAlaIleTyrThrGlyAspAspValIleArgIleIleIleAspAsp 208
DB 140 TGAACCG-AAAATACCACTTGGCGCTTACACAGGAGACAGACTCCCGCATATCCCGTGA 82
OY 208 pMetLeuCyValIaGlyAsnSerGlnArgAspSerCysIlySerGlyAspSerGlyGlyProLe 228

```

Db 81 CAAGCTGTGTGCCGGGAAAGCCAGAGGAGGAGGACTCGTCAAGAGGCGACTGTGGAGGCGCCCTG 22

Qy 228 vVal1CybLysValAsnGlyThr 235

Db 21 -GTCTGCAG-GTGAATGGCAC 2

RESULT 7

CD672689 642 bp mRNA linear EST 24-JUN-2003

LOCUS Fg15806.5', mRNA sequence.

DEFINITION CD672689

ACCESSION CD672689

VERSION CD672689.1 GI:32174420

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 642)

AUTHORS Mammalian, J.W., Bouffard, G., Smith, D. and Peterson, K. Touchman, J.W., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A., Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K. Expressed sequence tag analysis of adult human iris for the NEIBank project: steroid-response factors and similarities with retinal pigment epithelium

TITLE Mol. Vis. 8 (4), 185-195 (2002)

JOURNAL 22103462

MEDLINE 12107412

PUBMED 12107412

COMMENT Contact: Mistow G. Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 15 row: a column: 06

Seq primer: M13RP1 reverse primer (ABT).

FEATURES

source

1..642

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="Fg15806"

/tissue_type="Iris"

/dev_stage="Adult"

/lab_host="EMDHI0B"

/clone_lib="Human Iris cDNA (Normalized): fg"

/note="Organ: Eye; Vector: pCMVSPORT6; A human iris library (dx) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NcoI. This NcoI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Lite Technologies). Single-stranded DNA (1 mg) was hybridized (Cot 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.: 1,23e-95 Length: 642

Score: 1081.00 Matches: 198

Percent Similarity: 93.75% Conservative: 5

Best Local Similarity: 93.40% Mismatches: 9

Query Match: 71.49% Indels: 0

DB: 6 Gaps: 0

US-09-598-982C-52 (1-275) x CD672689 (1-642)

Qy 1 MetLeuSerLeuLeuLeuLeuLeuLeuValLeuProValLeuAlaSerArgAlaTyrAlaIpro 20

Db 6 ATGCTGAATCTGCTGCTGCTGGCCCTCCGCTGGGAGAGCCCGCTACCGCCGCCCT 65

Qy 21 AlAproValAlnAlaLeuGlnGlnAlaGlyTlleValIglVglVglnGlnAlaProArgSer 40

Db 66 GCCCAAGCCGAGCCCTGTGAGGAGTGGCACTGTGGAGGCTCAAGAGAGCCCAAGAGCC 125

Qy 41 LysTrpProTropGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60

Db 126 AAGTGGCCCTGTGCAAGTGTAGAGTCCGGCAGCCGATGTGAGTATGCATCTCTCGGG 185

Qy 61 GlySerLeuIleHisProGlnTTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80

Db 186 GGCTCCCTCATCCACCCCAAGTGGGTGGCAAGCCGCAAGCCGCTCTGCGTGGGACCCG 245

Qy 81 LysAspIleValAlaThrLeuArgValGlnLeuArgGlnHisIleuLeuTyrGlnAspGln 100

Db 246 AAGGATCTGGCCCGCTCAAGGCTGCAAGTGGGAGAGCACTCTACTACAGAGCAAG 305

Qy 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120

Db 306 CTGCTGGCCGTTAGAGAGATTCCTGCACCCACAGTTCCTACACGGCCAGATGGAGCG 365

Qy 121 AspIleAlaLeuLeuGlnLeuGlnGlnLeuProValAlaAlaSerSerArgValHisThrVal 140

Db 366 GACATCCGCCCTCTGAGAGTGGAGAGCCGGTGAACCTCCAGCCAGTCCACCGCTC 425

Qy 141 MetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrpValThrGlyTrrp 160

Db 426 GCCCTGCCCCCTGCTCAGAGACTTCCCGGGAGATCGGTGGGTGGTGG 485

Qy 161 GlYAspValAspAspGlnUProLeuProProPheProLeuUlyGlnValYAspVal 180

Db 486 GGCAGATGGCAAAAGATGAGGCGCTCCACCGCATTTCCCTGAAACAGGTGAGAGTTC 545

Qy 181 ProIleMetGluAsnHisIleCysAspAlaIlystrYrHisLeuGlyAlaTyrThrGlyAsp 200

Db 546 CCCAATTAATGAAAMAACCAATTTGTACCGCAAAATTCACCTTGGCCCTACACCGGAGAG 605

Qy 201 AspValArgIleIleIleArgAspMetLeuCysAla 212

Db 606 GAGCTCCGATCTCGTGAAGCAGATGCTGTGTC 641

RESULT 8

CF553058 806 bp mRNA linear EST 22-SEP-2003

LOCUS AGENCOURT_15594982 NIH_MGC_183 Homo sapiens cDNA clone

DEFINITION IMAGE:30529468.5', mRNA sequence.

ACCESSION CF553058

VERSION CF553058.1 GI:34889892

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 806)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

CONTACT Daniela S. Geithard, Ph.D.

OFFICE Office of Cancer Genomics / NIH

NATIONAL CANCER INSTITUTE

Bldg. 31 Rml0A07 Bethesda, MD 20892

EMAIL: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

PLACE: NDM6517 row: b column: 05

High quality sequence start: 18
 High quality sequence stop: 611.
 Location/Qualifiers
 1..806
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30529468"
 /lab_host="DH10B-TONa (T1 and T5 phage resistant)"
 /clone_lib="NIH_MGC_183"
 /note="Organ: Pooled muscle (cardiac and skeletal);
 Vector: PCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:
 NotI; Library is oligo-dT primed and directionally cloned
 (BcoRV site is destroyed upon cloning). Average insert
 size 1.7. Library was constructed by Invitrogen."

ORIGIN
 Alignment Scores:
 Pred. No.: 3,42e-93 Length: 806
 Score: 1057.50 Matches: 202
 Percent Similarity: 95.75% Conservative: 1
 Best Local Similarity: 95.28% Mismatches: 6
 Query Match: 69.94% Indels: 3
 DB: 7 Gaps: 1

US-09-598-982C-52 (1-275) x CF553058 (1-806)
 QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
 DB 158 ARGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 217
 QY 21 AlaProValGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
 DB 218 GCCCCAGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277
 QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
 DB 278 AAGTGGCCCTGGCAGGTGAGCTGAGAGAGTCCCGGACCGGATCTGATGATGATGATGATGATGATG 337
 QY 61 GlySerLeuLeuHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
 DB 338 GGCTCCCTCATCCACCCCGAGTGGGTGCTGACCGCGCGGCGCATGCTCCCTCCGAGCCGGACGCTC 397
 QY 81 LysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyrTyrGlnAspGln 100
 DB 398 AAGGATCTGGCCACTCGAGGTGCAACGCGGAGACGACCCCTTACTACAGAGACCGAG 457
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
 DB 458 CTGCTGCCCGGTGACGAGATCATCTCGTGCACCCACAGATTCTTACATCATCCAGACTGGAGCG 517
 QY 121 AspIleAlaLeuLeuGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
 DB 518 GATATATGCCCTCTGAGCTGGAGGAGGCCCGGTGAACATCTCCAGCCCGCTCCACACAGCGGTC 577
 QY 141 MetLeuProProAlaSerGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 160
 DB 578 ATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
 QY 161 GlyAspValAspAspAspGlnProLeuProProProPheProLeuGlyGlnValIleVal 180
 DB 638 GCGCGATGTGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 697
 QY 181 ProIleMetGlnAsnHisIleCysAspAlaLysTyrHisLeuGly--AlaTyrThrGlyA 200
 DB 698 CCCATATATGAAACCAACCACTTGTGTGACGCAAAATACACCTTGGGGCCCTTACACGGGAG 757
 QY 200 spAsp---ValArgIleIleIleArgAspAspMet 209
 DB 758 ACGAACCTCCGGGATCATCCCGTACGACATG 789

RESULT 9
 CB203717

LOCUS CB203717 930 bp mRNA linear EST 05-FEB-2003
 DEFINITION AGENCOURT 11288628 NIH MGC 135 Mus musculus cDNA clone
 IMAGE:30142408 5', mRNA sequence.
 ACCESSION CB203717
 VERSION CB203717.1 GI:28240351
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 TITLE
 AUTHORS
 REFERENCE
 COMMENT
 Email: cgarbs-remail.nih.gov
 Tissue Procurement: Dr. David Rowe
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDM0051 row: j column: 17
 High quality sequence stop: 681.
 Location/Qualifiers

FEATURES
 source
 1..930
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30142408"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_135"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Normalized full-length enriched library from pooled mouse
 embryonic limb, maxilla and mandible, day 12.5, 13.5,
 14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
 Cloned directionally, priming method: Oligo-dT. cDNA
 enrichment: >1x dp, Average insert size 1.6k dp.
 Normalization (Col value): 7.5 kb. Priming sequence:
 5'-GACTAGTCTTAGATCGCGGCGCCGCCCTG3' Tissue contributed
 by, David Rowe. Library constructed by Resgen, Invitrogen
 Corp."

ORIGIN
 Alignment Scores:
 Pred. No.: 8,82e-92 Length: 930
 Score: 1044.00 Matches: 189
 Percent Similarity: 82.76% Conservative: 27
 Best Local Similarity: 72.41% Mismatches: 39
 Query Match: 69.05% Indels: 6
 DB: 6 Gaps: 1

US-09-598-982C-52 (1-275) x CB203717 (1-930)
 QY 17 TyrAlaAlaProAlaProValGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 36
 DB 6 TACTTAGCCCTCCGCCCA-----GCCAATGAGGAGTGGCAVCTGGGAGGACATGAG 59
 QY 37 AlaProAspSerLysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMet 56
 DB 60 GCTTCTGAGATTAAGTGGCCCTTGGCAGGTGAGCTGAGATTAATAATTAATTAATTAATTAATTA 119
 QY 57 HisPheCysGlyGlySerLeuLeuHisProGlnTrpValLeuThrAlaAlaHisCysLeu 76
 DB 120 CATTTCTGCGGAGGCTCTCATCCACCCACAGTGGGTGCTCACTGGGGCCACACTGTGTG 179
 QY 77 GlyProAspValLysAspLeuAlaThrLeuArgValGlnLeuArgGlnGlnHisLeuTyr 96
 DB 180 GAAACCGACATCAAAAGCCCAAGCTCTTCCGGGTGCACTTCCGTGAGCAGTATCTATAC 239
 QY 97 TyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPheTyrIle 116

DB 240 TATGGGACCCAGCTCTCTTTTAAACCGGATCGTGGTGCACCCCCCACTATTACAGCGCC 299
 QY 117 GATThG1VA1AAspIleAlaLeuGluLeuGluGluUpProValAsnIleSerSerArg 136
 DB 300 GAGGATGGGGACAAGCCTTGGCTTGGAGCTTGGAGGTCCCTGTGAATGTCTCCACCCT 359
 QY 137 ValHisThrValMetLeuProProAlaSerGluThrPheProProGlyMetProCysTrrp 156
 DB 360 ATCCACCACCAATATCCCTGCCCCCTCGCCTCGGAGACCTTCCCTCCCGAGCAATCGTCTGG 419
 QY 157 ValThiGluYTrpG1VAAspValAspAsnAspGluProLeuProProPheProLeuLeu 176
 DB 420 GTAACAGGCTGGGGCCGCAATGATAAAGACGCCCTCTCCACCCTTAATCTTGAAG 479
 QY 177 GlnValIysValProIleMetGluAAsnHisIleCysAspAlaIalyIstYrHisLeuG1YAla 196
 DB 480 CAAGTGAAGTTCCTCCATTTGGGAAAACAAGCCCTGTGTGCACCGGAATACCACTAGCCCTTC 539
 QY 197 TyrThrG1VAAspValArG1leIleArgAspAspMetLeuGlyAsnSerGln 216
 DB 540 TACACGGGAGATGATTTTCCCATTTGCCATGATGGCAATGCTGTGTGTGGAAATACACAGG 599
 QY 217 ArgAspSerCysIysGlyAspSerGlyGlyProLeuValCysIysValAsnG1YTrpTrrp 236
 DB 600 AGAGACTCCCTGCCAGCGCGATTCAGGGGGCCCACTGGTCTGCAAAAGGTAAGTACCCTGG 659
 QY 237 LeuGlnAlaGlyValValIleSerTrp-AspGluGluCysAlaGlnProAsnArpProG1Yll 256
 DB 660 CTGCAGGCAAGAGATGTGTCACCTGGGAGTGGAGGTGGCAGACCCCAAGACCTGGCANT 719
 QY 256 eYrYrThrArg-ValIThr-TYrYrLeuAspTrrpIleHis-HisYrValProllys 273
 DB 720 CTACACCCGGGGTGAACATATNCTACTGATCAGTCCACCGCTATGCTTCGAA 774

RESULT 10
 CD671891
 LOCUS
 DEFINITION Eg08b10.y2 Human Iris cDNA (Normalized): Eg Homo sapiens cDNA clone
 ACCESSION CD671891
 VERSION CD671891.1 GI:32173622
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Taxonomy: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnindaey; Homo.
 REFERENCE 1 (bases 1 to 635)
 Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A., Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
 Expressed sequence tag analysis of adult human iris for the NEIBank Project: steroid-response factors and similarities with retinal pigment epithelium
 Mol. Vis. 8 (4), 185-195 (2002)
 JOURNAL MEDLINE
 PUBMED 22103462
 COMMENT
 Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 08 row: b column: 10
 Seg primer: M13RP1 reverse primer (AB1).
 Location/Qualifiers
 1..635
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fg08b10"
 /tissue_type="Iris"
 /dev_stage="Adult"
 /lab_host="EMDHI0B"

Alignment Scores:
 Pred. No.: 1 61e-91 Length: 635
 Score: 1039.00 Matches: 191
 Percent Similarity: 93.78% Conservative: 5
 Best Local Similarity: 91.39% Mismatches: 13
 Query Match: 68.72% Indels: 0
 DB: 6 Gaps: 0

US-09-598-982c-52 (1-275) x CD671891 (1-635)
 QY 1 MetLeuSerLeuLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaIyrAlaIaIaPro 20
 DB 7 ATGCTGAAATCTGCTGCTGCTGCTGCTGCTGCCCGTCCCTGGGAGCGCGCTCTACCGCCCTC 66
 QY 21 AlaprovalGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnAlaProArgSer 40
 DB 67 GCGCCAGCGCCAGCGCCCTGAGCGATGCGGCGGATCGTGGGGGGTCAAGAGAGCGCCCAAGAGC 126
 QY 41 LyrTrrpProTrrpGlnValSerIleuAlaArgAspArGlyTrrpMetHisPheCysGly 60
 DB 127 AAGTGGCCCTGGCAGGAGTGCCTGAGAGTCATAGGCCCAATAGTATGCACACTTCGCGGG 186
 QY 61 GlySerLeuIleHisProGlnTrrpValIleuThrAlaAlaHisCysLeuG1YProAspVal 80
 DB 187 GCCTCCCTCATTCACACCCCAATGGGAGTGCCTGCCGAGCGCACCTGGGAGCGGAGCACTC 246
 QY 81 LysAspLeuAlaThrIleuArgValGlnLeuArgGlnIleHisIleuYrYrGlnAspGln 100
 DB 247 AAGGATCTGGCCGCCCTCAGGGTGCACCTGCGGGAGCGACACTCTACTACGAGACCCAG 306
 QY 101 LeuLeuProValSerArgIleIleValHisProGlnPheYrIleIleG1YTrpG1YAla 120
 DB 307 CTGCTGCCCGGTAGAGAGATTCATGTCGACCCCGCACTTCAACCGCCCAAGTCCGAGCG 366
 QY 121 AspIleAlaLeuLeuGlnIleuGlnIleuProValAsnIleSerSerArgValHisThrVal 140
 DB 367 GACATGCCCTCTGAGTGGAGAGGCCGCGTGAACGTCCTCCAGCACAGTCCACACGCGGTC 426
 QY 141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysYrYrValIThrGlyTrrp 160
 DB 427 ACCCTGCCCCCTCGCTCAGAGACTTTCCCGGGGAAAGCCCGTGGTCTCTGCTGCTGG 486
 QY 161 GlYAspValAAspAsnAspGluProLeuProProPheProLeuLeuGlyValIlysVal 180
 DB 487 GGCGATGGAGAAATGATAGGCGCTCCACCGCCATTTCTTAAACAGAGTGAAGTTC 546
 QY 181 ProIleMetGluAAsnHisIleCysAspAlaIyIstYrHisIleuGlyAlaIyIstYrHisGlyAsp 200
 DB 547 CCAATAAATGAAAACCACTTTTGTAGCGAAATACCACTTTGGGGCTTACACAGGGAAGAC 606
 QY 201 AspValIarG1leIleArgAspAspMet 209
 DB 607 GACGTCGGCAGTCCTCGTAGACACAVG 633

RESULT 11
 BI413250

LOCUS B1413250 1000 bp mRNA linear EST 14-AUG-2001
 DEFINITION 602986390F1 NCI_GAP_LuJ3 Mus musculus cDNA clone IMAGE:5142478 5',
 ACCSSION B1413250
 VERSION B1413250.1 GI:15174173
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabds-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library: Arrayed by: The I.M.A.G.E. Consortium (ILNI)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNI at:
 http://image.llnl.gov
 plate: LLAM11351 row: c column: 23
 High quality sequence start: 32
 High quality sequence stop: 845.
 Location/Qualifiers

1.1000
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5142478"
 /issue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI CGAP LuJ3"
 /notes="Organ: lung; Vector: pTR73D-Pac (Pharmacia) with a
 modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAAGTGGAGCGGCCCTCTGTTTGTTTTGTGTGT
 TGTTACCAATCTGAAAGTGGAGCGGCCCTCTGTTTGTTTTGTGTGT
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pTR73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
 Alignment Scores:
 Pred. No.: 2.9e-90 Length: 1000
 Score: 1029.00 Matches: 189
 Percent Similarity: 81.274 Conservative: 28
 Best Local Similarity: 70.794 Mismatches: 43
 Query Match: 66.064 Indels: 7
 DB: 4 Gaps: 1

US-09-598-982c-52 (1-275) x B1413250 (1-1000)
 QY 11 ValLeuAlaSerArgAlaTyrAlaAlaProValAlaGlnAlaLeu-GlnGlnAlaG1 30
 Db 20 ATTATNCTTAAAGTCTGGTGTGACTCAAGCCCTCGCCCA-----GCCAAATCAAGCGAGTG66 73
 QY 30 y1LeVal1G1yG1nG1nAlaPrOArGSeRyLTPpPpOTpTg1nValSeriLeuArgVa 50
 Db 74 CATCCGTGGAGGACATAGAGCTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCTGAGATT 133
 QY 50 lArgdArpArgYyTTRMetHisPheCysG1yG1ySeriLeu1His1ProG1nTTPValle 70
 Db 134 TAAATTAACCTACTGGATTCATTTTGCGGAGGCTCTCATCCACCCACAGTGGTGCT 193
 QY 70 uPTrAlaAlaHisCysLeuG1yPrOArPaRyAllyLSaerLeuAlaThrLeuArgValGlnle 90

Db 194 CACTGCGGACACACTGTGTGGAGCCGACATCAAAAAGCCAGACTCTTCGCGGTGACCT 253
 QY 90 uArgGlnGlnHis1SreuYrYrGlnAspGlnLeuLeuProValSerArg1Le1leValHi 110
 Db 254 TCGTAGACGATATCAATCAATGAGGGACACACTCTCTTTTGAACCGGATGATGATGCA 313
 QY 110 pProGlnPheYr1Le1leGlnThrG1yAlaAsp1LeAlaLeuLeuGlnGlnG1nG1n 130
 Db 314 CCCCACATTTACACAGCCGAGGGTGGGACAGAGTTGCCCTGTGGAGCTTGGAGTCC 373
 QY 130 oValAsn1leSerSerArgValHisPthValMetLeuProProAlaSerG1uThrPhePr 150
 Db 374 TGTGAAATGTCCTCAACCATATTCACCCCATATCCCTGCGCCCTCGGAGACTTCC 433
 QY 150 oProG1yMetProCysTTPVal1ThrG1yTrng1yAspValAspAsnAspG1nProLeuPr 170
 Db 434 CCGTAGGACATCGTCTGGGAGACAGGCTGGGGGACATTAATTAATGACGAGCCCTCC 493
 QY 170 oProProPheProLeuLysGlnValLysValPro11eMetG1uAsnHis1leCyAspAsp1 190
 Db 494 ACCTCTTATCTCTGAAAGCAAGTGAAGTTCCTCATTTGTGGAAAACAGCCGTGTGACCG 553
 QY 190 aLys-TyrHisLeuG1yAlaTyrThrG1yAspAspValArg1Le1leArgAspAspMetL 210
 Db 554 GAAGTAAACACACATCGGCTCTACACGGGAGATGATTTTCATTCATGATGAGCATGCG 613
 QY 210 euCyAla1AG1yAsnSerGlnArgAspSerCyLysG1yAspSerG1yProLeuValC 230
 Db 614 TGTGTGCTGAAATVACGAGGAGACTCTCTCCAGGGGATTCAGGGGGGACACTGCT 673
 QY 230 ySLysValAsnG1yThrTrpLeuGlnAlaG1yValValSer-TTPAspG1nG1yCyAla 249
 Db 674 GCAAAGTGAAGAGGTRCCTGGCTGCAGGAGAGTGTGATCACCTGGGGTGAAGGGCTGCCA 733
 QY 250 Gln-ProAsnArgProG1yLe1yYrThrArgValThr-Tyr-TyLeuAspTTP1leHisH 269
 Db 734 CAGCCCAAAACAAGCTGGCATCTACACCCGGGTGACCATACTTACTTAACTGATCGATCCACG 793
 QY 269 lEYrYrValPro 272
 Db 794 CTTATGTCCT 804

RESULJ 12
 BUD071733
 LOCUS BUD071733 589 bp mRNA linear EST 27-AUG-2002
 DEFINITION hm33a09.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036761
 5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ; mRNA
 sequence.
 ACCSSION BUD071733
 VERSION BUD071733.1 GI:22512922
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 589)
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemshaja,I., Searce,M., Brestelli,V., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Birstein,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisshylli,R.,
 Williams,T., Jackson,Y. and Bowers,Y.
 TITLE Endocrine Pancreas Consortium
 JOURNAL Unpublished (2000)
 COMMENT Other ESTs: hm33a09.x1
 CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept Of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer *In vivo* mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu) Seq primer: -40RP from Gibco High quality sequence stop: 433.

FEATURES

Location/Qualifiers

1..589

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6036761"
 /tissue_type="Insulinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Human insulinoma"
 /note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZapII system (Stratagene) by Dr. J. Ferrer, *in vivo* mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

ORIGIN

Alignment Scores:

Pred. No.:	2,51e-88	Length:	589
Score:	1006.00	Matches:	180
Percent Similarity:	95.85%	Conservative:	5
Best Local Similarity:	93.26%	Mismatches:	8
Query Match:	66.53%	Indels:	0
DB:	5	Gaps:	0

US-09-598-982C-52 (1-275) x BU071733 (1-589)

```

QY 79 AspValIysAspLeuA1aThrLeuArgValGlnLeuArgGlnGlnHisLeuTYRTrgln 98
DB 10 GAGCGTAAAGATCTGGCGCCCTCAGGGTGCACCTGGGGAGCAGCACCCTTACTACCG 69
QY 99 AspGlnLeuLeuProValSerArgIleIleValHisProGlnPheTYRleIleGlnThr 118
DB 70 GACCACCTGTGCGCGGTGACAGATCATCTGCACCCACAGTTCTTACACCCGCCGAGATC 129
QY 119 G1VAlAspIleAlaLeuLeuGlnGlnGlnProValAsnIleSerArgValHis 138
DB 130 GAGCGGACATCCCTCTGCTGAGCTGGAGAGCCGGTGAACGTCCTCCAGCCGTCAC 189
QY 139 ThrValMetLeuProProAlaSerGlnThrPheProProGlyMetProCysTrpValThr 158
DB 190 ACGGTCAACCCCTGCTCCCTCAGAGACCTTCCCTCCCGGGGAGTGCCTGGTGGTCACT 249
QY 159 G1YTrpG1YAspValAspAsnAspGluProLeuProProPheProIleuLeuGlnVal 178
DB 250 GGTGGGGCCATGTGACATGTAGCGCCCTCCACCCGATTTCCCTTGAAGCGAGTG 309
QY 179 IysValProIleMetGluAsnHisIleCysAspAlaIleTYRHisIleGlnValAlaTYRThr 198
DB 310 AAGGTCCCAATAATGAAAACAATTTGTGACGCAAAATAACCACTTGGCCCTTACAG 369
QY 199 G1YAspAspValArgIleIleArgAspAspMetLeuCysAlaG1YAsnSerGlnArgAsp 218
DB 370 GAGACGACGATCCCGATCCGTCGTCAGACATGTCGTGTGCGGGAAACAACCCGGAGGAG 429
QY 219 SerCysIleG1YAspSerG1YIleProLeuValCysIleValAsnG1YThrTrpLeuGln 238
DB 430 TCAATGCCAGGGCCACTCCGGAGGGCCCTGTGTGCAAGGTGATGCACCTGGCTGCAG 489
QY 239 AlaG1YValIleValSerTrpAspG1YIleCysAlaGlnProAsnArgProG1YIleTYRThr 258
DB 490 GCGGGCGGTCACTGGGGCGAGGGGCTGTGCGCCAGCCCAACCGGCTGGCATTACACC 549

```

QY 259 ArgValIleThrTYRLeuAspTrpIleHisIleTYRVal 271
 DB 550 CGTGTCACTACTACTGTGAGCTGAGTCCACCACCACTATGTGC 588

RESULT 13
 BG697540
 LOCUS BG697540
 DEFINITION BG697540.1 GI:13963867
 mRNA sequence.

ACCESSION BG697540 811 bp mRNA linear EST 07-MAY-2001
 BG697540
 VERSION BG697540
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA
 Sequencing by: Inyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LHM10700 row: 9 column: 16
 High quality sequence stop: 781.
 Location/Qualifiers

FEATURES

source

1..811
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4803879"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Skn3"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NciI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:	1,63e-86	Length:	811
Score:	989.50	Matches:	200
Percent Similarity:	77.41%	Conservative:	9
Best Local Similarity:	74.07%	Mismatches:	17
Query Match:	65.44%	Indels:	45
DB:	4	Gaps:	3

US-09-598-982C-52 (1-275) x BG697540 (1-811)

```

QY 1 MetLeuSerLeuLeuLeuAlaLeuProValIleuAlaSerArgAlaTYRAlaIlePro 20
DB 11 AAGCTGAATTCGCTCTGCTGGCGCTGCCCGCTCCGCGGAGCGCGGCTTACAGCCGGC-CT 69
QY 21 AlaProValGlnAlaLeuGlnGlnAlaG1YIleValG1YGlnGlnIleAlaProArgSer 40
DB 70 GCGCCAGGCCAGGCCCTCAGGAGATGGCATTCGTTGGGGGTCCAGGAGCCGCCAGGAGC 129
QY 41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTYRTrpMetHisPheCysG1Y 60
DB 130 AAGTGCCTTGGCAAGTAACTTGAAGTCCAGGCCCAATCTGGATGCACTTTCTGGCGAG 189
QY 61 G1YSerLeuIleHisProGlnTrpValLeuThrAlaIleHisCysValG1YProAspVal 80
DB 190 GGTCCCTCATCCACCCCAAGTGGGTGTGAACCGGAGGCACTGGTGGAGCCGGAGCGTTC 249
QY 81 LysAspLeuAlaIleThrLeuArgValGlnLeuArgGlnHisIleValTYRTrpGlnAspGln 100

```

Db 250 AAGATCTGGCCGCTGAGGCTGACCTGGGAGGAGGACCTTACTACAGGACCG 309
 Qy 101 Leuleuprovalaserargllelevalhisproglinphetryllelelnthrglyala 120
 Db 310 CTGCTGCGCCGATGAGGATGATCGGACCCACAGATTCTACACCCGCGAGATCCGAGGCG 369
 Qy 121 Aapllealaleuleugluenlunlunluprovalasnliseserargvalhisrthval 140
 Db 370 GACATCGCCCTGCTGAGCTGGAGGAGCGGATGAAAGTCTCCAGCCACGTCACAGGCTC 429
 Qy 141 Metleuprovalaserglurthrpheprogllymetprocystrpvalthrlytrp 160
 Db 430 ACCCTGCCCCCTGCTCAGAGACCTTCCCCCGGGGATGCCCTGCTGGTCACTGGCTG 489
 Qy 161 Glyaarpvalaspaahsp----- 166
 Db 490 GGGATGATGACAAATGATGATGGTCTGGGGACAGTGGAGGTGGGCGCCAGGGTCTTAGCC 549
 Qy 166 ----- 166
 Db 550 ACAGCCCAAGCCCTGGGCTCCCTTGGGCTCCAGGTGGGGGTTGCCCGCCCTCCTGAG 609
 Qy 167 ---GluProLeuPro-----Pro-ProPheProLeuLysGlnValIlyeVa 180
 Db 610 GGTGACCTCTTCCACCTGACGAGAGGCTCCACGCACTTCTCTGAAAGCAGGTGAAGT 669
 Qy 180 lProilemetgluamhisileCyakspalalystrhis-leuglyalalythrlytrp 200
 Db 670 CCCCATAATGAAACCAACATTTGTGACGGCAAGATACACTTTGGCCGCTACAGGGAG 729
 Qy 200 spAaPvAlArGllelleaRgaPaSP-MetLeuCyAlaGlyAmSerGlnArGaSPeR 219
 Db 720 AGGACGTCGCGATCGCTCCGTCGACGAACATGCTGTGTGGGAAACCCGAG---GACTCA 786
 Qy 220 CyblysglyAaSPserGlyGlyPro 227
 Db 787 TCCCAAGGGGACTCCCGAGGGGCC 810

RESULT 14
 BU787772 581 bp mRNA linear EST 11-OCT-2002
 LOCUS i034c08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128247
 DEFINITION 5' similar to SW:TRYS_HUMAN P20231 BETA-TRYPYASB PRECURSOR ; mRNA
 sequence.
 accession BU787772.1 GI:23835713
 version BU787772
 keywords EST.
 source Homo sapiens (human)
 organism Homo sapiens
 reference Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 authors Eukaryota; Euteria; Primates; Carnivora; Homiidae; Homo.
 title Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,
 Lemliha, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marr, M., Pape, D., Wylie, F., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ralder, E., Ronko, I., Bennett, J.,
 Caridans, M., Gibbons, M., McCann, R., Cole, R., Tsagarelashvili, R.,
 Williams, T., Jackson, Y. and Bowers, Y.
 journal Endocrine Pancreas Consortium
 comment Unpublished (2000)
 other ESTs: i034c08.x1
 contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@bionh.harvard.edu
 Library was constructed by Dr. J. Ferrer in vivo mass-excised to
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
 University Genome Sequencing Center For information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@am.wustl.edu)
 Seg primer: -40RP from Gibco
 High quality sequence stop: 427.
 Location/Qualifiers
 1..581
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6128247"
 /issue_type="Insulinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="Human insulinoma"
 /note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
 pBluescript SK- by Dr. H. Inoue following the Washington
 University protocol
 (http://genome.wustl.edu/est/lambda_protocol.shtml).
 please contact Hiroshi Inoue, MD/PhD for further
 information on this library (Metabolism division, Permut
 Laboratory, Washington University School of Medicine, Box
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
 is a Washington University Pancreas EST project library."

FEATURES
 source
 Alignment Scores:
 Pred. No.: 1.01e-83 Length: 581
 Score: 959.00 Matches: 176
 Percent Similarity: 93.72% Conservative: 3
 Best Local Similarity: 92.15% Mismatches: 12
 Query Match: 63.43% Indels: 0
 DB: Gaps: 0

US-09-598-982c-52 (1-275) x BU787772 (1-581)

Qy 5 LeuleuAlaleuAlaleuAlaleuAlaserArgAlaRyAlaAlaProAlaProValGln 24
 Db 8 CTGCTGCTGGGCGTCCGCTCTGCTGCGGAGCGCGGCTTACGGGGCCCTGCGCCAGGCGAG 67

Qy 25 AlaleuGlnGlnAlaGlyIleValGlyGlnGlnAlaProArGserlyeTrpTrp 44
 Db 68 CCCCCTGACGACGAGGGGATGCTGGGGGATGAGAGGGCCCGCAGGAGCACTGGCGCTG 127

Qy 45 GlnValSerLeuAryValAryAspaRgyrTrpMetHisPheCyGlySerLeuIle 64
 Db 128 CAGGTGACCTGAGAGTCCACGCGCCCACTGAGTGCCTTCTGGGGGCTCCCTCATC 187

Qy 65 HisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValIlyeAspLeuAla 84
 Db 188 CACCCCGAGTGGTGTCTGACCGGACGCGACTGCGTGGAGCCGAGACGTCAAGATCTGGCC 247

Qy 85 ThrLeuArGValGlnLeuAryGlnHisLeuVryTyGlnAerGlnLeuProVal 104
 Db 248 GCCCTCAGGGTGGCAACTCGGGGAGCAGCACTTACTTACAGGACCAAGCTGCTGGCGTTC 307

Qy 105 SerArglleleValHisPheProglinphetryllelelnthrglyalaAapllealaleu 124
 Db 308 AGCAGGATCACTGGACCCACAGTTCACACCCGCGCATGCGAGCGGAATCGCCCTG 367

Qy 125 LeuGluLeuGluGluProValAsnIleSerSerArgValHisrthValMetLeuProPro 144
 Db 368 CTGAGAGCTGGAGAGCCGGTGAAGTCTCCAGCCACGTCACAGGTCACCTGCCCTC 427

Qy 145 AlaSerGlnurthrpheprogllymetprocystrpvalthrlytrpGlyAaPvAlaSP 164
 Db 428 GCTTCAAGAGACCTTCCCCCGGGGATGCGTGGTCACTGGCTGGGGGAGATGTGGAC 487

Qy 165 AaPaSPGluProLeuProProPheProLeuLysGlnValIlyeValProIleMetGly 184
 Db 488 AATGATGAGCCCTCCACCGCACTTCTCTGAAAGCGGTGAAAGTCCCAATAATGAA 547

Qy 185 AaPnHisIleCyAspAlaLalyeTyTrHisleuGly 195

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:29:16 ; Search time 114 Seconds
(without alignments)
947.848 Million cell updates/sec

Title: US-09-598-982C-52
Perfect score: 1512
Sequence: 1 MSLILLALPVLASRAVAAP.....ITTRVTVYLDIMHVVPKKP 275

Scoring table: BLOSSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues
Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
21:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1512	100.0	275	14	US-10-117-323-36 Sequence 36, App1
2	1512	100.0	275	15	US-10-352-684A-46 Sequence 46, App1
3	1414	93.3	275	15	US-10-287-226-90 Sequence 92, App1
4	1411	93.3	275	15	US-10-287-226-92 Sequence 92, App1
5	1411	93.3	275	16	US-10-723-860-2690 Sequence 2690, App1
6	1411	93.3	275	18	US-10-756-149-5451 Sequence 5451, App1
7	1405	92.9	275	10	US-09-813-432-44 Sequence 44, App1
8	1405	92.9	275	15	US-10-352-684A-48 Sequence 48, App1
9	1405	92.9	275	15	US-10-174-364-44 Sequence 44, App1
10	1405	92.9	275	15	US-10-246-583-44 Sequence 44, App1
11	1405	92.9	275	15	US-10-287-226-94 Sequence 94, App1

ALIGNMENTS

Query	DB	Score	DB ID	Length	Indels	Gaps
1	MSLILLALPVLASRAVAAPVQALDQAGIVGQEARSRKMPVQVSLRVRDRYMMHFCG 60	100.0%	US-10-117-323-36	275	0	0
1	MSLILLALPVLASRAVAAPVQALDQAGIVGQEARSRKMPVQVSLRVRDRYMMHFCG 60	100.0%	US-10-117-323-36	275	0	0

RESULT 1
US-10-117-323-36 Application US/10117323
; Sequence 36, Application US/10117323
; Publication No. US20030054992A1
; GENERAL INFORMATION:
; APPLICANT: Rancourt, Derrick E.
; APPLICANT: Rancourt, Susan L.
; TITLE OF INVENTION: Implantation Sertine Proteinases
; FILE REFERENCE: 033337-005
; CURRENT FILING DATE: 2002-04-08
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/294,736
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/350,962
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-323-36

Query Match 100.0%; Score 1512; DB 14; Length 275;
Best Local Similarity 100.0%; Pred. No. 4.5e-130;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GSLIHQWVLTAAHCLGSPVDKDLATLRVQLREOHLYYDDQLLPSRRIIVHPQFYIIQTGA 120
 DB 61 GSLIHQWVLTAAHCLGSPVDKDLATLRVQLREOHLYYDDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRHVHTVMTLPPASETFPPGMPQWVTGMDVNDDEPLPPFPPLKQYKV 180
 DB 121 DIALLELEBPVNISSRHVHTVMTLPPASETFPPGMPQWVTGMDVNDDEPLPPFPPLKQYKV 180
 QY 181 PIMENHICDAKHYLGAAYTGDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWTLOAG 240
 DB 181 PIMENHICDAKHYLGAAYTGDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWTLOAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHVVYPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHVVYPKKP 275

RESULT 2
 US-10-352-684A-46

; Sequence 46, Application US/10352684A
 ; Publication No. US20030215452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals Inc.
 ; APPLICANT: Carroll, Joseph M.
 ; APPLICANT: Healy, Aileen
 ; APPLICANT: Kelly, Nadine S.
 ; APPLICANT: Kelly, Louise M.
 ; APPLICANT: Weich, Nadine S.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
 ; TITLE OF INVENTION: 1513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
 ; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
 ; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
 ; FILE REFERENCE: MP102-019PRNOMNIM
 ; CURRENT APPLICATION NUMBER: US/10/352,684A
 ; CURRENT FILING DATE: 2003-01-28
 ; PRIOR APPLICATION NUMBER: US 60/354,333
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: US 60/360,258
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/364,476
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: US 60/375,626
 ; PRIOR FILING DATE: 2002-04-26
 ; PRIOR APPLICATION NUMBER: US 60/386,494
 ; PRIOR FILING DATE: 2002-06-06
 ; PRIOR APPLICATION NUMBER: US 60/390,965
 ; PRIOR FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: US 60/392,480
 ; PRIOR FILING DATE: 2002-06-28
 ; PRIOR APPLICATION NUMBER: US 60/394,128
 ; PRIOR FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: US 60/399,783
 ; PRIOR FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: US 60/403,221
 ; PRIOR FILING DATE: 2002-08-13
 ; Remaining Prior Application data removed - See file wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 275
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-352-684A-46

Query Match 100.0%; Score 1512; DB 15; Length 275;
 Best Local Similarity 100.0%; Pred. No. 4,5e-130;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAAYAAPVQALQAGIVGQEPAPSSKMPWQVSLRVRDRYMMHFGG 60
 DB 1 MSLILLALPVLASRAAYAAPVQALQAGIVGQEPAPSSKMPWQVSLRVRDRYMMHFGG 60
 QY 61 GSLIHQWVLTAAHCLGSPVDKDLATLRVQLREOHLYYDDQLLPSRRIIVHPQFYIIQTGA 120

DB 61 GSLIHQWVLTAAHCLGSPVDKDLATLRVQLREOHLYYDDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRHVHTVMTLPPASETFPPGMPQWVTGMDVNDDEPLPPFPPLKQYKV 180
 DB 121 DIALLELEBPVNISSRHVHTVMTLPPASETFPPGMPQWVTGMDVNDDEPLPPFPPLKQYKV 180
 QY 181 PIMENHICDAKHYLGAAYTGDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWTLOAG 240
 DB 181 PIMENHICDAKHYLGAAYTGDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWTLOAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHVVYPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHVVYPKKP 275

RESULT 3
 US-10-287-226-90

; Sequence 90, Application US/10287226
 ; Publication No. US20040086875A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agee, Michele L.,
 ; APPLICANT: Alsobrook, John P.,
 ; APPLICANT: Berghs, Constance,
 ; APPLICANT: Boldog, Ference,
 ; APPLICANT: Burgess, Catherine E.,
 ; APPLICANT: Chant, John S.,
 ; APPLICANT: Chaudhuri, Amitabha,
 ; APPLICANT: Dipippo, Vincent A.,
 ; APPLICANT: Edinger, Shlomit R.,
 ; APPLICANT: Eisen, Andrew,
 ; APPLICANT: Ellerman, Karen,
 ; APPLICANT: Gangolli, Esha A.,
 ; APPLICANT: Gorman, Linda,
 ; APPLICANT: Gerlach, Valerie,
 ; APPLICANT: Ji, Weizhen,
 ; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Khramtsov, Nikolai,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Malyanakar, Uriel M.,
 ; APPLICANT: MacDougall, John R.,
 ; APPLICANT: Mezes, Peter S. E.,
 ; APPLICANT: Miller, Charles E.,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Ooi, Chean Eng,
 ; APPLICANT: Ort, Tatiana,
 ; APPLICANT: Padigara, Muralidhara,
 ; APPLICANT: Pattarajan, Meera,
 ; APPLICANT: Rastelli, Luca,
 ; APPLICANT: Rieger, Daniel K.,
 ; APPLICANT: Rothenberg, Mark E.,
 ; APPLICANT: Shenoy, Suresh G.,
 ; APPLICANT: Spaderna, Steven K.,
 ; APPLICANT: Spyrek, Kimberley A.,
 ; APPLICANT: Taudier, Jr., Raymond J.,
 ; APPLICANT: Vernet, Corine A.M.,
 ; APPLICANT: Zernusen, Bryan D.,
 ; APPLICANT: Zhong, Mei
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-480C
 ; CURRENT APPLICATION NUMBER: US/10/287,226
 ; CURRENT FILING DATE: 2002-11-04
 ; PRIOR APPLICATION NUMBER: 2001-11-30
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: 60/334,421
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/354,392
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/360,148
 ; PRIOR FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: 60/364,000
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 60/404,821
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: 60/334,526

? PRIOR FILING DATE: 2001-11-30
 ? PRIOR APPLICATION NUMBER: 60/354,409
 ? PRIOR FILING DATE: 2002-02-04
 ? PRIOR APPLICATION NUMBER: 60/364,227
 ? PRIOR FILING DATE: 2002-03-13
 ? PRIOR APPLICATION NUMBER: 60/334,027
 ? PRIOR FILING DATE: 2001-11-28
 ? PRIOR APPLICATION NUMBER: 60/331,641
 ? PRIOR FILING DATE: 2001-11-20
 ? Remaining Prior Application data removed - See File Wrapper or PALM.
 ? NUMBER OF SEQ ID NOS: 673
 ? SOFTWARE: Cursesqlst version 0.1
 ? SEQ ID NO 90
 ? LENGTH: 275
 ? TYPE: PRT
 ? ORGANISM: Homo sapiens
 ? US-10-287-226-90

Query Match 93.5%; Score 1414; DB 15; Length 275;
 Best Local Similarity 92.7%; Pred. No. 4.3e-121;
 Matches 255; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAVAAPVQALQAGIVGGQEA PRSKMPQVSLRVDRYMMHFCG 60
 1 MSLILLALPVLASRAVAAPVQALQAGIVGGQEA PRSKMPQVSLRVDRYMMHFCG 60
 DB 61 GSLIHPOWVLTAAHCVBPDIKDLALRYQLREOHLYYODQLLPVSRITVHPQFYIOTGA 120
 61 GSLIHPOWVLTAAHCVBPDIKDLALRYQLREOHLYYODQLLPVSRITVHPQFYIOTGA 120
 QY 121 DIALLEEBEPVNISSRVHTVWLPASSETPPGMPGVCWTGWGVNDNDPPLPPFPKQVKV 180
 121 DIALLEEBEPVNISSRVHTVWLPASSETPPGMPGVCWTGWGVNDNDPPLPPFPKQVKV 180
 DB 181 PIMENHICDAKTYHGAAYTGDVRIIRDDMLCAGNSQSDSCGSGPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKTYHGAAYTGDVRIIRDDMLCAGNSQSDSCGSGPLVCKVNGTWTLOAG 240
 QY 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275
 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275
 DB 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275

RESULT 4
 US-10-287-226-92
 ? Sequence 92, Application US/10287226
 ? Publication No. US20040086875A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Agee, Michele L.
 ? APPLICANT: Alsobrook, John P.
 ? APPLICANT: Berghs, Constance,
 ? APPLICANT: Boldog, Ference,
 ? APPLICANT: Burgess, Catherine E.,
 ? APPLICANT: Chant, John S.,
 ? APPLICANT: Chaudhuri, Amitabha,
 ? APPLICANT: DiPippo, Vincent A.,
 ? APPLICANT: Edinger, Shlomit R.,
 ? APPLICANT: Eissen, Andrew,
 ? APPLICANT: Ellerman, Karen,
 ? APPLICANT: Gangolli, Bsha A.,
 ? APPLICANT: Gorman, Linda,
 ? APPLICANT: Gerlach, Valerie,
 ? APPLICANT: Ji, Weizhen,
 ? APPLICANT: Kekuda, Ramesh,
 ? APPLICANT: Khrantsev, Nikolai,
 ? APPLICANT: Li, Li,
 ? APPLICANT: Malyanekar, Uriel M.,
 ? APPLICANT: MacDougall, John R.,
 ? APPLICANT: Mezes, Peter S.,
 ? APPLICANT: Miller, Charles E.,
 ? APPLICANT: Millet, Isabelle,
 ? APPLICANT: Ooi, Chuan Eng,
 ? APPLICANT: Ort, Tatiana,

? APPLICANT: Padigaru, Muralidhara,
 ? APPLICANT: Patturajan, Meera,
 ? APPLICANT: Patelli, Luca,
 ? APPLICANT: Rieger, Daniel K.,
 ? APPLICANT: Rothenberg, Mark E.,
 ? APPLICANT: Shenoy, Suresh G.,
 ? APPLICANT: Spaderna, Steven K.,
 ? APPLICANT: Spytek, Kimberley A.,
 ? APPLICANT: Taupier, Jr., Raymond J.,
 ? APPLICANT: Venet, Corine A.M.,
 ? APPLICANT: Zernusen, Bryan D.,
 ? TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ? FILE REFERENCE: 21402-480C
 ? CURRENT APPLICATION NUMBER: US/10/287,226
 ? PRIOR FILING DATE: 2002-11-04
 ? PRIOR APPLICATION NUMBER: 60/334,421
 ? PRIOR FILING DATE: 2001-11-30
 ? PRIOR APPLICATION NUMBER: 60/354,392
 ? PRIOR FILING DATE: 2002-02-04
 ? PRIOR APPLICATION NUMBER: 60/360,148
 ? PRIOR FILING DATE: 2002-02-27
 ? PRIOR APPLICATION NUMBER: 60/364,000
 ? PRIOR FILING DATE: 2002-03-13
 ? PRIOR APPLICATION NUMBER: 60/404,821
 ? PRIOR FILING DATE: 2002-08-30
 ? PRIOR APPLICATION NUMBER: 60/334,526
 ? PRIOR FILING DATE: 2001-11-30
 ? PRIOR APPLICATION NUMBER: 60/354,409
 ? PRIOR FILING DATE: 2002-02-04
 ? PRIOR APPLICATION NUMBER: 60/364,227
 ? PRIOR FILING DATE: 2002-03-13
 ? PRIOR APPLICATION NUMBER: 60/334,027
 ? PRIOR FILING DATE: 2001-11-28
 ? PRIOR APPLICATION NUMBER: 60/331,641
 ? Remaining Prior Application data removed - See File Wrapper or PALM.
 ? NUMBER OF SEQ ID NOS: 673
 ? SOFTWARE: Cursesqlst version 0.1
 ? SEQ ID NO 92
 ? LENGTH: 275
 ? TYPE: PRT
 ? ORGANISM: Homo sapiens
 ? US-10-287-226-92

Query Match 93.3%; Score 1411; DB 15; Length 275;
 Best Local Similarity 92.4%; Pred. No. 8.1e-121;
 Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MSLILLALPVLASRAVAAPVQALQAGIVGGQEA PRSKMPQVSLRVDRYMMHFCG 60
 1 MSLILLALPVLASRAVAAPVQALQAGIVGGQEA PRSKMPQVSLRVDRYMMHFCG 60
 DB 61 GSLIHPOWVLTAAHCVBPDIKDLALRYQLREOHLYYODQLLPVSRITVHPQFYIOTGA 120
 61 GSLIHPOWVLTAAHCVBPDIKDLALRYQLREOHLYYODQLLPVSRITVHPQFYIOTGA 120
 QY 121 DIALLEEBEPVNISSRVHTVWLPASSETPPGMPGVCWTGWGVNDNDPPLPPFPKQVKV 180
 121 DIALLEEBEPVNISSRVHTVWLPASSETPPGMPGVCWTGWGVNDNDPPLPPFPKQVKV 180
 DB 181 PIMENHICDAKTYHGAAYTGDVRIIRDDMLCAGNSQSDSCGSGPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKTYHGAAYTGDVRIIRDDMLCAGNSQSDSCGSGPLVCKVNGTWTLOAG 240
 QY 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275
 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275
 DB 241 VVSMDEGCAQPNRPFGITRVTYTYLDMIHVYVPKKP 275

RESULT 5
 US-10-723-860-2690
 ? Sequence 2690, Application US/10723860

```

; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2690
; LENGTH: 275
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-723-860-2690

```

Query Match 93.3%; Score 1411; DB 16; Length 275;

Best Local Similarity 92.4%; Pred. No. 8,1e-121;

Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

```

QY 1 MSLILLIAPVLAASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60
DB 1 MNLMLLILALPVLASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60
QY 61 GSLIHPOWVLTAAHCIGPVDKDLATLRVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120
DB 61 GSLIHPOWVLTAAHCIGPVDKDLALRLVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120
QY 121 DIALLELEBPVNVSSRVHTVMLPPASSETPPGMPQWVTGMDVNDDEPLRPPPLKQVKY 180
DB 121 DIALLELEBPVNVSSRVHTVTLRPPASSETPPGMPQWVTGMDVNDDEPLRPPPLKQVKY 180
QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQDSCGDSGGLVCKVNGTMDQAG 240
DB 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNTRRSDSCGDSGGLVCKVNGTMDQAG 240
QY 241 VVSMDEGCAQPNRPGIYTRVTVYLDWTHHYVPKKP 275
DB 241 VVSMDEGCAQPNRPGIYTRVTVYLDWTHHYVPKKP 275

```

RESULT 6

```

; Sequence 5451, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5451
; LENGTH: 275
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-756-149-5451

```

Query Match 93.3%; Score 1411; DB 18; Length 275;

Best Local Similarity 92.4%; Pred. No. 8,1e-121;

Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

```

QY 1 MSLILLIAPVLAASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60
DB 1 MNLMLLILALPVLASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60

```

```

QY 61 GSLIHPOWVLTAAHCIGPVDKDLATLRVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120
DB 61 GSLIHPOWVLTAAHCIGPVDKDLALRLVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120
QY 121 DIALLELEBPVNVSSRVHTVMLPPASSETPPGMPQWVTGMDVNDDEPLRPPPLKQVKY 180
DB 121 DIALLELEBPVNVSSRVHTVTLRPPASSETPPGMPQWVTGMDVNDDEPLRPPPLKQVKY 180
QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQDSCGDSGGLVCKVNGTMDQAG 240
DB 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNTRRSDSCGDSGGLVCKVNGTMDQAG 240
QY 241 VVSMDEGCAQPNRPGIYTRVTVYLDWTHHYVPKKP 275
DB 241 VVSMDEGCAQPNRPGIYTRVTVYLDWTHHYVPKKP 275

```

RESULT 7

```

; Sequence 44, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Jr., Raymond J
; APPLICANT: Majumder, Kamud
; APPLICANT: Spaderma, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vermet, Corine A. M.
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 275
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-813-432-44

```

Query Match 92.9%; Score 1405; DB 10; Length 275;

Best Local Similarity 92.0%; Pred. No. 2,9e-120;

Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

```

QY 1 MSLILLIAPVLAASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60
DB 1 MNLMLLILALPVLASRAVAAPVQALQOAGIVGGQEARPRSKMPQVSLRVDRYMMHFCG 60
QY 61 GSLIHPOWVLTAAHCIGPVDKDLATLRVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120
DB 61 GSLIHPOWVLTAAHCIGPVDKDLALRLVQLREQHLYYODQLLPSRRIIVHPQFYIIQTGA 120

```

QY 121 DIALLELEBPVNISSRVTWMLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPPLKQVKV 180
 DB 121 DIALLELEBPVNISSRVTWMLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPPLKQVKV 180
 QY 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNSQRDSCGSDGSGPLVCKVNGTWMLOAG 240
 DB 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNTRRSDSCGSDGSGPLVCKVNGTWMLOAG 240
 QY 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275
 DB 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275

RESULT 8

US-10-352-684A-48
 / Sequence 48, Application US/10352684A
 / Publication No. US20030215452A1
 / GENERAL INFORMATION:
 / APPLICANT: Millennium Pharmaceuticals Inc.
 / APPLICANT: Carroll, Joseph M.
 / APPLICANT: Healy, Aileen
 / APPLICANT: Welch, Nadine S.
 / APPLICANT: Kelly, Louise M.
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 / TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
 / TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28058, 7366, 12212,
 / TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
 / TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
 / FILE REFERENCE: MP102-019PRNOMNTM
 / CURRENT APPLICATION NUMBER: US/10/352,684A
 / CURRENT FILING DATE: 2003-01-28
 / PRIOR APPLICATION NUMBER: US 60/354,333
 / PRIOR FILING DATE: 2002-02-04
 / PRIOR APPLICATION NUMBER: US 60/360,258
 / PRIOR FILING DATE: 2002-02-28
 / PRIOR APPLICATION NUMBER: US 60/364,476
 / PRIOR FILING DATE: 2002-03-15
 / PRIOR APPLICATION NUMBER: US 60/375,626
 / PRIOR FILING DATE: 2002-04-26
 / PRIOR APPLICATION NUMBER: US 60/386,494
 / PRIOR FILING DATE: 2002-06-06
 / PRIOR APPLICATION NUMBER: US 60/390,965
 / PRIOR FILING DATE: 2002-06-24
 / PRIOR APPLICATION NUMBER: US 60/392,480
 / PRIOR FILING DATE: 2002-06-28
 / PRIOR APPLICATION NUMBER: US 60/394,128
 / PRIOR FILING DATE: 2002-07-03
 / PRIOR APPLICATION NUMBER: US 60/399,783
 / PRIOR FILING DATE: 2002-07-31
 / PRIOR APPLICATION NUMBER: US 60/403,221
 / Remaining Prior Application data removed - See file Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 62
 / SOFTWARE: Patent Ver. 2.1
 / SOFTWARE: Faetseg for Windows Version 4.0
 / SEQ ID NO 48
 / LENGTH: 275
 / TYPE: PRT
 / ORGANISM: Homo Sapiens
 / US-10-352-684A-48

Query Match 92.9%; Score 1405; DB 15; Length 275;
 Best Local Similarity 92.0%; Pred. No. 2.9e-120;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLLLLALPVLASRAVAAPAVQALQAGIVGGQEARSKMPQVSLRVDRYMMHFCG 60
 DB 1 MNTLLLLALPVLASRAVAAPAVQALQAGIVGGQEARSKMPQVSLRVDRYMMHFCG 60
 QY 61 GSLIHPQWVLTAAKCLGPRVVDLALRVQLREOHLYYDQDLLPVSRIIVHPQFYITQGA 120
 DB 61 GSLIHPQWVLTAAKCLGPRVVDLALRVQLREOHLYYDQDLLPVSRIIVHPQFYITQGA 120

QY 121 DIALLELEBPVNISSRVTWMLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPPLKQVKV 180
 DB 121 DIALLELEBPVNISSRVTWMLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPPLKQVKV 180
 QY 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNSQRDSCGSDGSGPLVCKVNGTWMLOAG 240
 DB 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNTRRSDSCGSDGSGPLVCKVNGTWMLOAG 240
 QY 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275
 DB 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275

RESULT 9

US-10-174-364-44
 / Sequence 44, Application US/10174364
 / Publication No. US20030216308A1
 / GENERAL INFORMATION:
 / APPLICANT: Anderson et al.
 / FILE REFERENCE: 15966-729C1P2
 / TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
 / CURRENT APPLICATION NUMBER: US/10/174,364
 / CURRENT FILING DATE: 2002-06-17
 / PRIOR APPLICATION NUMBER: 60/190,835
 / PRIOR FILING DATE: 2000-03-20
 / PRIOR APPLICATION NUMBER: 60/190,768
 / PRIOR FILING DATE: 2000-03-20
 / PRIOR APPLICATION NUMBER: 60/190,972
 / PRIOR FILING DATE: 2000-03-22
 / PRIOR APPLICATION NUMBER: 60/191,199
 / PRIOR FILING DATE: 2000-03-22
 / PRIOR APPLICATION NUMBER: 60/191,947
 / PRIOR FILING DATE: 2000-03-24
 / PRIOR APPLICATION NUMBER: 60/192,665
 / PRIOR FILING DATE: 2000-03-28
 / PRIOR APPLICATION NUMBER: 60/192,657
 / PRIOR FILING DATE: 2000-03-28
 / PRIOR APPLICATION NUMBER: 60/192,836
 / PRIOR FILING DATE: 2000-03-29
 / Remaining Prior Application data removed - See file Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 128
 / SOFTWARE: Patent Ver. 2.1
 / SOFTWARE: Patent Ver. 2.1
 / SEQ ID NO 44
 / LENGTH: 275
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-10-174-364-44

Query Match 92.9%; Score 1405; DB 15; Length 275;
 Best Local Similarity 92.0%; Pred. No. 2.9e-120;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLLLLALPVLASRAVAAPAVQALQAGIVGGQEARSKMPQVSLRVDRYMMHFCG 60
 DB 1 MNTLLLLALPVLASRAVAAPAVQALQAGIVGGQEARSKMPQVSLRVDRYMMHFCG 60
 QY 61 GSLIHPQWVLTAAKCLGPRVVDLALRVQLREOHLYYDQDLLPVSRIIVHPQFYITQGA 120
 DB 61 GSLIHPQWVLTAAKCLGPRVVDLALRVQLREOHLYYDQDLLPVSRIIVHPQFYITQGA 120
 QY 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNSQRDSCGSDGSGPLVCKVNGTWMLOAG 240
 DB 181 PIMENHICDAXKHLGAYTGDDVRIIRDDMLCAGNTRRSDSCGSDGSGPLVCKVNGTWMLOAG 240
 QY 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275
 DB 241 VVSMDEGCAQENRPGIYTRVYTYLDMVHHVYPKKP 275

Db 241 VVSMGEGCAQPNRRPGIYTRVYLYLDMWHYVPPKKP 275

RESULT 10
US-10-246-583-44
; Sequence 44, Application US/10246583
; Publication No. US2004058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-583-44

Query Match 92.9%; Score 1405; DB 15; Length 275;
Best Local Similarity 92.0%; Pred. No. 2.9e-120;
Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;
QY 1 MSLLLLALPVLASRAVAAPVQALQQAAGIVGGGAPRSPKMPWVSLRVRDRYMMHFGG 60
Db 1 MNLVLLALPVLASRAVAAPVQALQQAAGIVGGGAPRSPKMPWVSLRVRDRYMMHFGG 60
QY 61 GSLIHPQWVLTAAHCLGPRVDKDLATLRVQLREOHLYYDQDLRPFVSR11VHPQFY11QTGA 120
Db 61 GSLIHPQWVLTAAHCLGPRVDKDLATLRVQLREOHLYYDQDLRPFVSR11VHPQFY11QTGA 120
QY 121 DIALLLBERPVNISSRVHTYMLPPASSTPPRPGMPCWVTGKGDVNDDEPLRPPRPLKQVRY 180
Db 121 DIALLLBERPVNISSRVHTYMLPPASSTPPRPGMPCWVTGKGDVNDDEPLRPPRPLKQVRY 180
QY 181 PIMENHICDAKHYKLGAVTGDVRIIRDDMLCAGNSORDSCGKSGGPIVCKVNGTWLQAG 240
Db 181 PIMENHICDAKHYKLGAVTGDVRIIRDDMLCAGNSORDSCGKSGGPIVCKVNGTWLQAG 240
QY 241 VVSMGEGCAQPNRRPGIYTRVYLYLDMWHYVPPKKP 275
Db 241 VVSMGEGCAQPNRRPGIYTRVYLYLDMWHYVPPKKP 275

RESULT 11
US-10-287-226-94
; Sequence 94, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:

APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Bergns, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eisen, Andrew,
APPLICANT: Ellemann, Karen,
APPLICANT: Gangolli, Esha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khrantsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: MaYankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Orc, Tatiana,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Patutrajau, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spylek, Kimberley A.,
APPLICANT: Taubler, Jr., Raymond J.,
APPLICANT: Vernet, Corine A.,
APPLICANT: Zehusen, Bryan D.,
APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 94
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-94
Query Match 92.9%; Score 1405; DB 15; Length 275;
Best Local Similarity 92.0%; Pred. No. 2.9e-120;
Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLIILALPVLASRAVAAPAPVQALQOAGIVGGEAPRSKMPQVSLRVDRYMMHFCG 60
 DB 1 MNLILILALPVLASRAVAAPAPGQALQVGIWGGGAPRSKMPQVSLRVDRYMMHFCG 60
 QY 61 GSLIHPQWVLTAAHCLGPDVVDLTLRVQLREOHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
 DB 61 GSLIHPQWVLTAAHCLGPDVVDLTLRVQLREOHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
 QY 121 DIALLLEBPVNISRRVHTVTLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPKQVKV 180
 DB 121 DIALLLEBPVNISRRVHTVTLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPKQVKV 180
 QY 181 PIMENHICDAKYHLGAYTGDVRIIRDMLCAGNSQDSCCKDSSGGPLVCKVNGTWTLOAG 240
 DB 181 PIMENHICDAKYHLGAYTGDVRIIRDMLCAGNSQDSCCKDSSGGPLVCKVNGTWTLOAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHYYVPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHYYVPKKP 275

RESULT 12

US-10-689-832-44
 ; Sequence 44, Application US/10689832
 ; Publication No. US20040121380A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder, Kamud
 ; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
 ; FILE REFERENCE: 15966-729DIV1
 ; CURRENT APPLICATION NUMBER: US/10/689, 832
 ; FILE REFERENCE: PT-0085 USN
 ; CURRENT APPLICATION NUMBER: US/10/275, 505
 ; PRIOR FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: 60/190, 835
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190, 768
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190, 972
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191, 199
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191, 947
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 60/192, 665
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192, 657
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192, 984
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192, 664
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 44
 ; LENGTH: 275
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-689-832-44

Query Match 92.9%; Score 1405; DB 16; Length 275;
 Best Local Similarity 92.0%; Pred. No. 2.9e-120;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 121 DIALLLEBPVNISRRVHTVTLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPKQVKV 180
 DB 121 DIALLLEBPVNISRRVHTVTLPPASSTFPFGMPCWVTGMDVNDDEPLPPFPKQVKV 180
 QY 181 PIMENHICDAKYHLGAYTGDVRIIRDMLCAGNSQDSCCKDSSGGPLVCKVNGTWTLOAG 240
 DB 181 PIMENHICDAKYHLGAYTGDVRIIRDMLCAGNSQDSCCKDSSGGPLVCKVNGTWTLOAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHYYVPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVYTYLDMIHYYVPKKP 275

RESULT 13

US-10-275-505-13
 ; Sequence 13, Application US/10275505
 ; Publication No. US20040081961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: DELBERGANE, Angelo M.; LAL, Preeti G.
 ; APPLICANT: HARALITA, April J.A.; PATTERSON, Chandra
 ; APPLICANT: WALITA, Nandier K.; KEARNEY, Liam
 ; APPLICANT: TRIBOUTIEY, Catherine M.; KHAN, Farrah A.
 ; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
 ; APPLICANT: AZIMZAI, Yalda; ELIHOTT, Yicki S.
 ; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
 ; APPLICANT: YANG, Junning; HERNANDEZ, Roberto
 ; APPLICANT: POLICKI, Jennifer L.; LU, Dyoung Aina M.
 ; APPLICANT: REDDY, Roopa M.; YDE, Henry
 ; APPLICANT: TANG, Y. Tom
 ; TITLE OF INVENTION: PROTEASES
 ; FILE REFERENCE: PT-0085 USN
 ; CURRENT APPLICATION NUMBER: US/10/275, 505
 ; CURRENT FILING DATE: 2002-11-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/14651
 ; PRIOR FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/209, 402
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: 60/207, 477
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: 60/205, 803
 ; PRIOR FILING DATE: 2000-05-17
 ; PRIOR APPLICATION NUMBER: 60/203, 566
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: 60/202, 082
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PERL Program
 ; SEQ ID NO: 13
 ; LENGTH: 691
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 7472460CD1
 ; US-10-275-505-13

Query Match 92.7%; Score 1401; DB 15; Length 691;
 Best Local Similarity 92.0%; Pred. No. 2e-119;
 Matches 253; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

QY 181 PIMENHICDAKYHLGAYTGDVRIIRDDMICAGNSQRDSCGKDSGGPILVCKVNGTWMLOAG 240
 DB 189 PIMENHICDAKYHLGAYTGDVRIIRDDMICAGNTRRSDSCGKDSGGPILVCKVNGTWMLOAG 248
 QY 241 VVSMDEGCAQPNRPGIYTRVYLLDMTHHVPKPP 275
 DB 249 VVSMDEGCAQPNRPGIYTRVYLLDMTHHVPKPP 283

RESULT 14
 US-10-117-323-40
 ; Sequence 40, Application US//10117323
 ; Publication No. US20030054993A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rancourt, Derrick E.
 ; APPLICANT: Rancourt, Susan L.
 ; APPLICANT: O'Sullivan, Colleen M.
 ; TITLE OF INVENTION: Implantation Serine Proteinases
 ; FILE REFERENCE: 033337-005
 ; CURRENT APPLICATION NUMBER: US/10/117,323
 ; CURRENT FILING DATE: 2002-04-08
 ; PRIOR APPLICATION NUMBER: US 60/281,724
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: US 60/294,736
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: US 60/350,962
 ; PRIOR FILING DATE: 2002-01-25
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 40
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Mouse
 ; US-10-117-323-40

Query Match 91.7%; Score 1386.5; DB 14; Length 272;
 Best Local Similarity 91.9%; Pred. No. 1.4e-118;
 Matches 251; Conservative 8; Mismatches 13; Indels 1; Gaps 1;
 QY 3 SLLLLALPVLASRAVYAAAPAVVQALQQAGIVGGQEARPSKMPQVSLRVRDRYWMHFCGSS 62
 DB 1 NULLLLALPVLASRAVYAAAPAVVQALQGRVIVGGQEARPSKMPQVSLRVRDRYWMHFCGSS 60
 QY 63 LHPQVWVLAHCLGPDVQKDLATLRLVQLRBOHLIYDQDLIPVSRRIIVHPQFYIIOTGADI 122
 DB 61 LHPQVWVLAHCLGPDVQKDLATLRLVQLRBOHLIYDQDLIPVSRRIIVHPQFYIIOTGADI 119
 QY 123 ALLELEPVNISRVTWMLPASETFPPGMPQWVTKGWGDVNDDEPLPPEPLKQVKVPI 182
 DB 120 ALLELEPVNISRVTWMLPASETFPPGMPQWVTKGWGDVNDDEPLPPEPLKQVKVPI 179
 QY 183 MENHICDAKYHLGAYTGDVRIIRDDMICAGNSQRDSCGKDSGGPILVCKVNGTWMLOAGV 242
 DB 180 MENHICDAKYHLGAYTGDVRIIRDDMICAGNTRRSDSCGKDSGGPILVCKVNGTWMLOAGV 239
 QY 243 SMDEGCAQPNRPGIYTRVYLLDMTHHVPKPP 275
 DB 240 SMDEGCAQPNRPGIYTRVYLLDMTHHVPKPP 272

RESULT 15
 US-09-789-210-69
 ; Sequence 69, Application US//09789210
 ; Publication No. US20040241646A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COHEN, MAURICE
 ; APPLICANT: COBITTS, TRACEY L.
 ; APPLICANT: FRIEDMAN, PAULA N.
 ; APPLICANT: GRANADOS, EDWARD N.
 ; APPLICANT: KLAS, MICHAEL R.
 ; APPLICANT: RUSSELL, JOHN C.
 ; APPLICANT: STEWART, KENT D.
 ; APPLICANT: STROUP, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
 AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
 OF THE PROSTATE
 NUMBER OF SEQUENCES: 76
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Abbott Laboratories
 STREET: 100 Abbott Park Road
 CITY: Abbott Park
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/789,210
 FILING DATE: 20-Feb-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/944,483
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Becker, Cheryl L.
 REGISTRATION NUMBER: 35,441
 REFERENCE/DOCKET NUMBER: 6183.US.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847/935-1729
 TELEFAX: 847/938-2623
 TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 69:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 245 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: None
 SEQUENCE DESCRIPTION: SEQ ID NO: 69:
 US-09-789-210-69

Query Match 90.8%; Score 1373; DB 11; Length 245;
 Best Local Similarity 100.0%; Pred. No. 2.1e-117;
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 IVGGQEARPSKMPQVSLRVRDRYWMHFCGSSLHPQVWVLAHCLGPDVQKDLATLRLVQL 90
 DB 1 IVGGQEARPSKMPQVSLRVRDRYWMHFCGSSLHPQVWVLAHCLGPDVQKDLATLRLVQL 60
 QY 91 RBOHLIYDQDLIPVSRRIIVHPQFYIIOTGADIALLELEPVNISRVTWMLPASETFPP 150
 DB 61 RBOHLIYDQDLIPVSRRIIVHPQFYIIOTGADIALLELEPVNISRVTWMLPASETFPP 120
 QY 151 PGMPCWVTKGWGDVNDDEPLPPEPLKQVKVPI MENHICDAKYHLGAYTGDVRIIRDDMIC 210
 DB 121 PGMPCWVTKGWGDVNDDEPLPPEPLKQVKVPI MENHICDAKYHLGAYTGDVRIIRDDMIC 180
 QY 211 CAGNSQRDSCGKDSGGPILVCKVNGTWMLOAGV VVSMDEGCAQPNRPGIYTRVYLLDMTHH 270
 DB 181 CAGNSQRDSCGKDSGGPILVCKVNGTWMLOAGV VVSMDEGCAQPNRPGIYTRVYLLDMTHH 240
 QY 271 VPKPP 275
 DB 241 VPKPP 245

Search completed: August 27, 2005, 23:41:06
 Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:22:30 ; Search time 63 Seconds
(without alignments)
325,849 Million cell updates/sec

Title: US-09-598-982C-52

Perfect score: 1512

Sequence: 1 MLSLLLLALPLVASRAVAAP.....ITRVTYVLDVIMHHVPRKP 275

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTOS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1512	100.0	275	2	US-09-016-366A-17	Sequence 17, Appl
2	1512	100.0	275	2	US-08-978-404B-12	Sequence 12, Appl
3	1402	92.7	273	2	US-09-016-366A-19	Sequence 19, Appl
4	1402	92.7	273	2	US-08-978-404B-14	Sequence 14, Appl
5	1400	92.6	274	2	US-09-016-366A-21	Sequence 21, Appl
6	1400	92.6	274	2	US-08-978-404B-16	Sequence 16, Appl
7	1394	92.2	267	2	US-09-016-366A-23	Sequence 23, Appl
8	1394	92.2	267	2	US-08-978-404B-18	Sequence 18, Appl
9	1394	92.2	267	4	US-09-917-254-101	Sequence 101, App
10	1373	90.8	245	3	US-08-944-483-69	Sequence 69, Appl
11	1284	84.9	245	3	US-09-079-970A-6	Sequence 6, Appl
12	1284	84.9	245	4	US-09-601-318-1	Sequence 1, Appl
13	1284	84.9	249	4	US-09-079-970A-5	Sequence 5, Appl
14	1277	84.5	244	4	US-09-601-318-4	Sequence 4, Appl
15	1277	84.5	244	4	US-09-601-318-5	Sequence 5, Appl
16	1277	84.5	244	4	US-09-601-318-6	Sequence 6, Appl
17	1277	84.5	244	4	US-09-601-318-7	Sequence 7, Appl
18	1134	75.0	273	2	US-08-978-404B-3	Sequence 3, Appl
19	1133	74.9	276	2	US-09-016-366A-15	Sequence 15, Appl
20	1133	74.9	276	2	US-08-978-404B-21	Sequence 21, Appl
21	1109.5	73.4	270	2	US-08-978-404B-8	Sequence 8, Appl
22	1105	73.1	273	2	US-08-978-404B-6	Sequence 6, Appl
23	1105	73.1	274	2	US-08-978-404B-5	Sequence 5, Appl
24	820	54.2	190	3	US-08-845-998-4	Sequence 4, Appl
25	820	54.2	190	3	US-09-206-537-4	Sequence 4, Appl
26	820	54.2	190	3	US-09-430-854-4	Sequence 4, Appl
27	816	54.0	190	2	US-08-845-998-6	Sequence 6, Appl

28	816	54.0	190	3	US-09-206-537-6	Sequence 6, Appl
29	816	54.0	190	2	US-09-430-854-4	Sequence 6, Appl
30	566.5	45.4	269	2	US-08-978-404B-10	Sequence 10, Appl
31	572.5	37.9	290	4	US-02-386-653A-7	Sequence 7, Appl
32	555	36.7	315	4	US-09-386-653A-9	Sequence 9, Appl
33	553.5	36.6	284	4	US-09-387-375-7	Sequence 7, Appl
34	553.5	36.6	284	4	US-10-041-400A-7	Sequence 7, Appl
35	553.5	36.6	284	4	US-10-042-091A-7	Sequence 7, Appl
36	535	35.4	316	4	US-09-387-375-9	Sequence 9, Appl
37	535	35.4	316	4	US-10-041-400A-9	Sequence 9, Appl
38	535	35.4	316	4	US-10-042-091A-9	Sequence 9, Appl
39	530.5	35.1	317	4	US-09-386-629-7	Sequence 7, App
40	530.5	35.1	317	4	US-09-907-794A-263	Sequence 263, App
41	530.5	35.1	317	4	US-09-905-125A-263	Sequence 263, App
42	530.5	35.1	317	4	US-09-902-775A-263	Sequence 263, App
43	530.5	35.1	317	4	US-09-906-700-263	Sequence 263, App
44	530.5	35.1	317	4	US-09-903-603A-263	Sequence 263, App
45	530.5	35.1	317	4	US-09-904-920A-263	Sequence 263, App

ALIGNMENTS

RESULT 1
US-09-016-366A-17
Sequence 17, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESS: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-016-366A-17
Query Match 100.0%; Score 1512; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 8,3e-153;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 60
 DB 1 MSLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 60
 QY 61 GSLIHPQWVLTAAHCLGPRVQKDLATLRVQLREOHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 DB 61 GSLIHPQWVLTAAHCLGPRVQKDLATLRVQLREOHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCWVTGWDVNDDEPLPPEPFLKQVKV 180
 DB 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCWVTGWDVNDDEPLPPEPFLKQVKV 180
 QY 181 PIMENHICDAKYHLGAAYTGDDVRIIRDMLCAGNSQRDSCKDGSGPLVCKVNGTWLQAG 240
 DB 181 PIMENHICDAKYHLGAAYTGDDVRIIRDMLCAGNSQRDSCKDGSGPLVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMIHYYVPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMIHYYVPKKP 275

RESULT 2

US-08-978-404B-12
 ; Sequence 12, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 275 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: NO. 5968782e
 ; US-08-978-404B-12

Query Match 100.0%; Score 1512; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 8.3e-153;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 60
 DB 1 MSLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 60

QY 61 GSLIHPQWVLTAAHCLGPRVQKDLATLRVQLREOHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 DB 61 GSLIHPQWVLTAAHCLGPRVQKDLATLRVQLREOHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCWVTGWDVNDDEPLPPEPFLKQVKV 180
 DB 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCWVTGWDVNDDEPLPPEPFLKQVKV 180
 QY 181 PIMENHICDAKYHLGAAYTGDDVRIIRDMLCAGNSQRDSCKDGSGPLVCKVNGTWLQAG 240
 DB 181 PIMENHICDAKYHLGAAYTGDDVRIIRDMLCAGNSQRDSCKDGSGPLVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMIHYYVPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVTVYLLDMIHYYVPKKP 275

RESULT 3

US-09-016-366A-19
 ; Sequence 19, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Huang, Chifu
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,366A
 ; FILING DATE: January 30, 1998
 ; CLASSIFICATION: 530
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/037,090
 ; FILING DATE: 05-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7093
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 273 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-016-366A-19

Query Match 92.7%; Score 1402; DB 2; Length 273;
 Best Local Similarity 92.3%; Pred. No. 4.4e-141;
 Matches 252; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 3 SLLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 62
 DB 1 NLLLLALPVLASRAVAAPAVQALQAGIVGGGEAPRSKMPWQVSLRVRDRYMMHFCG 60
 QY 63 LIHPQWVLTAAHCLGPRVQKDLATLRVQLREOHLYYQDQLLPSRRIIVHPQFYIIQTGADI 122

Db 61 LIHPQWVLTAAHCVGPDVKDLAALRVLREQHLYYQDQLPVSRIIHHVHPPFTYAQIGADI 120
 Qy 123 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 182
 Db 121 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 180
 Qy 183 MENNICAKYHLGATYGDVRIIRDVMLCAGNSORDSCCKDSSGGPVLCKVNGTMLQAGV 242
 Db 181 MENNICAKYHLGATYGDVRIIRDVMLCAGNTRDSCCKDSSGGPVLCKVNGTMLQAGV 240
 Qy 243 SWDEGCAQPNRPGIYTRVYTYLDMWHHYVPKKP 275
 Db 241 SWDEGCAQPNRPGIYTRVYTYLDMWHHYVPKKP 273

RESULT 4
 US-08-978-404B-14
 ; Sequence 14, Application US/08978404B
 ; Patent No. 5968782

GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESS: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978, 404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 273 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: NO. 5968782
 US-08-978-404B-14

Query Match 92.7%; Score 1402; DB 2; Length 273;
 Best Local Similarity 92.3%; Pred. No. 4.4e-141;
 Matches 252; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 3 SILLLALPVLASRAVAAPAPVQALQOAGIVGGQAPRSKMPQVSLRVRDRYMMHFCGG 62
 Db 1 NLLLLALPVLASRAVAAPAPVQALQOAGIVGGQAPRSKMPQVSLRVRDRYMMHFCGG 60
 Qy 63 LIHPQWVLTAAHCVGPDVKDLAALRVLREQHLYYQDQLPVSRIIHHVHPPFTYAQIGADI 122
 Db 61 LIHPQWVLTAAHCVGPDVKDLAALRVLREQHLYYQDQLPVSRIIHHVHPPFTYAQIGADI 120
 Qy 123 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 182

Db 121 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 180
 Qy 183 MENNICAKYHLGATYGDVRIIRDVMLCAGNSORDSCCKDSSGGPVLCKVNGTMLQAGV 242
 Db 181 MENNICAKYHLGATYGDVRIIRDVMLCAGNTRDSCCKDSSGGPVLCKVNGTMLQAGV 240
 Qy 243 SWDEGCAQPNRPGIYTRVYTYLDMWHHYVPKKP 275
 Db 241 SWDEGCAQPNRPGIYTRVYTYLDMWHHYVPKKP 273

RESULT 5
 US-09-016-366A-21
 ; Sequence 21, Application US/09016366A
 ; Patent No. 5953431

GENERAL INFORMATION:
 APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 NUMBER OF SEQUENCES: 65
 CORRESPONDENCE ADDRESS:
 ADDRESS: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,366A
 FILING DATE: January 30, 1998
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,090
 FILING DATE: 05-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7093
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 274 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-016-366A-21

Query Match 92.6%; Score 1400; DB 2; Length 274;
 Best Local Similarity 92.0%; Pred. No. 7.2e-141;
 Matches 252; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 2 SILLLALPVLASRAVAAPAPVQALQOAGIVGGQAPRSKMPQVSLRVRDRYMMHFCGG 61
 Db 1 NLLLLALPVLASRAVAAPAPVQALQOAGIVGGQAPRSKMPQVSLRVRDRYMMHFCGG 60
 Qy 62 SLIHPQWVLTAAHCVGPDVKDLAALRVLREQHLYYQDQLPVSRIIHHVHPPFTYAQIGADI 121
 Db 61 SLIHPQWVLTAAHCVGPDVKDLAALRVLREQHLYYQDQLPVSRIIHHVHPPFTYAQIGADI 120
 Qy 122 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 181
 Db 121 ALLEBEEVNISSRHTVMTLPPASSTFPPGMPVMTGMDVNDDEPLRPPPLKQVYPI 180

QY 182 IMENHI CDAKYHLGAYTGDDVRIIRDMLCAGNSORDSCKGSGGGLVCKVNGTWLQAGV 241
 DB 181 IMENHI CDAKYHLGAYTGDDVRIIRDMLCAGNTRRSDSCGDSGGGLVCKVNGTWLQAGV 240
 QY 242 VSMDEGCAQPNRPGIYTRVTVYLLDWMIHVYVPKP 275
 DB 241 VSMDEGCAQPNRPGIYTRVTVYLLDWMIHVYVPKP 274

RESULT 6

US-08-978-404B-16
 ; Sequence 16, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978, 404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354
 ; FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; INDEX:
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 274 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: No. 5968782e
 ; US-08-978-404B-16

Query Match 92.6%; Score 1400; DB 2; Length 274;
 Best Local Similarity 92.0%; Pred. No. 7.2e-141;
 Matches 252; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 2 LSLILLALPLVLSRAVAAPAPVQALQAGIVGGQEARPSKMPWQVSLRVRDRYMMHFFCGG 61
 DB 1 LNLMLLALPLVLSRAVAAPAPVQALQAGIVGGQEARPSKMPWQVSLRVRDRYMMHFFCGG 60
 QY 62 SLIHPQWVLTAAHCLGPDVVDLTLTRVQLREOHLYYQDQLLPVSRITIVHPQFYIIQTGAD 121
 DB 61 SLIHPQWVLTAAHCLGPDVVDLTLTRVQLREOHLYYQDQLLPVSRITIVHPQFYIIQTGAD 120
 QY 122 IALLLEBEFVNYSRVTHTVWMLPPASSTPPGMPCWVTGMDVNDDEPLPPPPPLKQVKVP 181
 DB 121 IALLLEBEFVNYSRVTHTVWMLPPASSTPPGMPCWVTGMDVNDDEPLPPPPPLKQVKVP 180
 QY 182 IMENHI CDAKYHLGAYTGDDVRIIRDMLCAGNSORDSCKGSGGGLVCKVNGTWLQAGV 241
 DB 181 IMENHI CDAKYHLGAYTGDDVRIIRDMLCAGNTRRSDSCGDSGGGLVCKVNGTWLQAGV 240

QY 242 VSMDEGCAQPNRPGIYTRVTVYLLDWMIHVYVPKP 275
 DB 241 VSMDEGCAQPNRPGIYTRVTVYLLDWMIHVYVPKP 274

RESULT 7

US-09-016-366A-23
 ; Sequence 23, Application US/09016366A
 ; Patent No. 5955431
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Huang, Chifu
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
 ; TITLE OF INVENTION: INHIBITORS
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,366A
 ; FILING DATE: January 30, 1998
 ; CLASSIFICATION: 530
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/037,090
 ; FILING DATE: 05-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7093
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; INDEX:
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-016-366A-23

Query Match 92.2%; Score 1394; DB 2; Length 267;
 Best Local Similarity 93.3%; Pred. No. 3e-140;
 Matches 249; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 9 LPLVLSRAVAAPAPVQALQAGIVGGQEARPSKMPWQVSLRVRDRYMMHFFCGGSLIHPQW 68
 DB 1 LPLVLSRAVAAPAPVQALQAGIVGGQEARPSKMPWQVSLRVRDRYMMHFFCGGSLIHPQW 60
 QY 69 VITAAHCLGPDVVDLTLTRVQLREOHLYYQDQLLPVSRITIVHPQFYIIQTGADLALLE 128
 DB 61 VITAAHCLGPDVVDLTLTRVQLREOHLYYQDQLLPVSRITIVHPQFYIIQTGADLALLE 120
 QY 129 EPNVYSSRVHTVWMLPPASSTPPGMPCWVTGMDVNDDEPLPPPPPLKQVKVPIMENHI 188
 DB 121 EPNVYSSRVHTVWMLPPASSTPPGMPCWVTGMDVNDDEPLPPPPPLKQVKVPIMENHI 180
 QY 189 DAKYHLGAYTGDDVRIIRDMLCAGNSORDSCKGSGGGLVCKVNGTWLQAGVSWDEGC 248
 DB 181 DAKYHLGAYTGDDVRIIRDMLCAGNTRRSDSCGDSGGGLVCKVNGTWLQAGVSWDEGC 240
 QY 249 AQPNRPGIYTRVTVYLLDWMIHVYVPKP 275

DB 241 AQPNRPGIYTRVTVYLDWIHHYVPPKP 267

RESULT 8

US-08-978-404B-18

Sequence 18, Application US/08978404B
Patent No. 5968782
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESS: Wolf, Greenfield & Sacks, P.C.
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-2441
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-18

Query Match 92.2%; Score 1394; DB 2; Length 267;
Best Local Similarity 93.3%; Pred. No. 3e-140;
Matches 249; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

9 LPVLAASRAVAAPAPVQALQOAGIVGGQEARPSKMPWQVSLRVRDRYMMHFCGSSLIHPQW 68
1 LPVLAASRAVAAPAPVQALQOAGIVGGQEARPSKMPWQVSLRVRDRYMMHFCGSSLIHPQW 60
69 VITAAHCLGPDVVDLATTREKRLVQREHLYYQDQLLPVSRRIIVHPQFYIIQTGADIALLELE 128
61 VITAAHCLGPDVVDLATTREKRLVQREHLYYQDQLLPVSRRIIVHPQFYIIQTGADIALLELE 120
129 EAVNISSRVHTVMLPPASSTFPPGMPGCVTGMGDVNDDEPLRPPPLKQVPIEMNHIC 188
121 EAVKVVSHVHTVTLPPASSTFPPGMPGCVTGMGDVNDDEPLRPPPLKQVPIEMNHIC 180
189 DAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKQDGGPVLCKVNGTMTLQAGVSWDGGC 248
181 DAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSQDSDGGPVLCKVNGTMTLQAGVSWDGGC 240
249 AQPNRPGIYTRVTVYLDWIHHYVPPKP 275
241 AQPNRPGIYTRVTVYLDWIHHYVPPKP 267

RESULT 9

US-09-917-254-101

Sequence 101, Application US/09917254
Patent No. 6703204
GENERAL INFORMATION:
APPLICANT: Muller, George
APPLICANT: Baak, Jan
TITLE OF INVENTION: Prognostic Classification of Breast Cancer
FILE REFERENCE: B0801/7224(JRV)
CURRENT APPLICATION NUMBER: US/09/917,254
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: US 60/222,093
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.0
SEQ ID NO 101
LENGTH: 267
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-917-254-101

Query Match 92.2%; Score 1394; DB 4; Length 267;
Best Local Similarity 93.3%; Pred. No. 3e-140;
Matches 249; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

9 LPVLAASRAVAAPAPVQALQOAGIVGGQEARPSKMPWQVSLRVRDRYMMHFCGSSLIHPQW 68
1 LPVLAASRAVAAPAPVQALQOAGIVGGQEARPSKMPWQVSLRVRDRYMMHFCGSSLIHPQW 60
69 VITAAHCLGPDVVDLATTREKRLVQREHLYYQDQLLPVSRRIIVHPQFYIIQTGADIALLELE 128
61 VITAAHCLGPDVVDLATTREKRLVQREHLYYQDQLLPVSRRIIVHPQFYIIQTGADIALLELE 120
129 EAVNISSRVHTVMLPPASSTFPPGMPGCVTGMGDVNDDEPLRPPPLKQVPIEMNHIC 188
121 EAVKVVSHVHTVTLPPASSTFPPGMPGCVTGMGDVNDDEPLRPPPLKQVPIEMNHIC 180
189 DAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKQDGGPVLCKVNGTMTLQAGVSWDGGC 248
181 DAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSQDSDGGPVLCKVNGTMTLQAGVSWDGGC 240
249 AQPNRPGIYTRVTVYLDWIHHYVPPKP 275
241 AQPNRPGIYTRVTVYLDWIHHYVPPKP 267

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
;
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 62324566
;
; US-08-944-483-69
;
; Query Match          90.8%; Score 1373; DB 3; Length 245;
; Best Local Similarity 100.0%; Pred. No. 4,6e-138;
; Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 31 IYGGGEAPRSKMPQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVVDLALRVQL 90
;   |||||
;   1 IYGGGEAPRSKMPQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVVDLALRVQL 60
; QY 91 REQHLYYQDQLLPSVRIIVHPQFYIIQTGADIALLEBEPVNISSRVHTVWMLPPASETFP 150
;   |||||
;   61 REQHLYYQDQLLPSVRIIVHPQFYIIQTGADIALLEBEPVNISSRVHTVWMLPPASETFP 120
; QY 151 PGMPCWVTGWGVDVNDDEPLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIIRDML 210
;   |||||
;   121 PGMPCWVTGWGVDVNDDEPLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIIRDML 180
; QY 211 CAGNSQRDSCKDGSGGSPLYCKVNGTMLQAGVYVSWDGGCAQPNRPRTYTRVYYLDMIHXY 270
;   |||||
;   181 CAGNSQRDSCKDGSGGSPLYCKVNGTMLQAGVYVSWDGGCAQPNRPRTYTRVYYLDMIHXY 240
; QY 271 VPKKP 275
;   |||||
;   241 VPKKP 245
;
; RESULT 11
; US-09-079-970A-6
; Sequence 6, Application US/09079970A
; Patent No. 6274366
; GENERAL INFORMATION:
; APPLICANT: Mafilit, Mark A.
; APPLICANT: Niles, Andrew L.
; APPLICANT: Haak-Frendscho, Mary
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human
; TITLE OF INVENTION: Beta-Trypsase and Method of Making Same
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Intellectual Property Department
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WISCONSIN
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,970A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leone, Joseph T.
; REGISTRATION NUMBER: 37,170
; REFERENCE/DOCKET NUMBER: 34506.073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-079-970A-6
;
; Query Match          84.9%; Score 1284; DB 3; Length 245;
; Best Local Similarity 92.7%; Pred. No. 1,4e-128;
; Matches 227; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
;
; QY 31 IYGGGEAPRSKMPQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVVDLALRVQL 90
;   |||||
;   1 IYGGGEAPRSKMPQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVVDLALRVQL 60
; QY 91 REQHLYYQDQLLPSVRIIVHPQFYIIQTGADIALLEBEPVNISSRVHTVWMLPPASETFP 150
;   |||||
;   61 REQHLYYQDQLLPSVRIIVHPQFYIIQTGADIALLEBEPVNISSRVHTVWMLPPASETFP 120
; QY 151 PGMPCWVTGWGVDVNDDEPLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIIRDML 210
;   |||||
;   121 PGMPCWVTGWGVDVNDDEPLPPFPPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIIRDML 180
; QY 211 CAGNSQRDSCKDGSGGSPLYCKVNGTMLQAGVYVSWDGGCAQPNRPRTYTRVYYLDMIHXY 270
;   |||||
;   181 CAGNTRRDSQDSSGSPLYCKVNGTMLQAGVYVSWDGGCAQPNRPRTYTRVYYLDMIHXY 240
; QY 271 VPKKP 275
;   |||||
;   241 VPKKP 245
;
; RESULT 12
; US-09-601-318-1
; Sequence 1, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft z. Prd. d. Wissensch. e.V
; APPLICANT: Byk Gulden Lomberg Chemische Fabrik GmbH
; APPLICANT: Bode, Wolfram
; APPLICANT: Moroder, Luis
; APPLICANT: Pereira, Pedro Jose Barbosa
; APPLICANT: Bergner, Andreas
; APPLICANT: Huber, Robert
; APPLICANT: Sommerhoff, Christian
; APPLICANT: Schaschke, No. 6613769bert
; APPLICANT: Br. Thomas
; APPLICANT: Martin, Thomas
; APPLICANT: Stadlwiesser, Josef
; APPLICANT: Ulrich, Wolf-Rüdiger
; APPLICANT: Dominik, Andreas
; APPLICANT: Thibaut, Ulrich
; APPLICANT: Bundschuh, Daniela
; APPLICANT: Beume, Rolf
; APPLICANT: Goebel, Karl-Josef
; TITLE OF INVENTION: TRYPTASE-INHIBITOREN
; FILE REFERENCE: 17674P WO-1
; CURRENT APPLICATION NUMBER: US/09/601,318
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: DE19804761.4

```

PRIOR FILING DATE: 1998-02-06
 PRIOR APPLICATION NUMBER: DE19851300.3
 PRIOR FILING DATE: 1998-11-06
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.1.1
 SEQ ID NO 1
 LENGTH: 245
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-601-318-1

Query Match 84.9%; Score 1284; DB 4; Length 245;
 Best Local Similarity 92.7%; Pred. No. 1.4e-128;
 Matches 227; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 31 IYGGQEARSKMPQVSLRVRDRYMMHFCGSSLIHPQWVTLAAHCLGPDVYKDLATLRVQL 90
 DB 1 IYGGQEARSKMPQVSLRVRDRYMMHFCGSSLIHPQWVTLAAHCLGPDVYKDLATLRVQL 60
 QY 91 RQHLIYQDOLLPSRRIIVHPQFYIIOTGADIALLEBEPNVISRRVHTWMLPPASETFP 150
 DB 61 RQHLIYQDOLLPSRRIIVHPQFYIIOTGADIALLEBEPNVISRRVHTWMLPPASETFP 120
 QY 151 PGMPCWVTGMGDVNDDEPLPPFPPLKQVVPIMENHICDAKYHLGAYTGDVRIIRDML 210
 DB 121 PGMPCWVTGMGDVNDDEPLPPFPPLKQVVPIMENHICDAKYHLGAYTGDVRIIRDML 180
 QY 211 CAGNSQRDSCKDSGGGLVCKVNGTWTLAGVYVSWDEGCAQPNRPGIYTRVYIYDMTHY 270
 DB 181 CAGNTRRDS CGDGGGLVCKVNGTWTLAGVYVSWDEGCAQPNRPGIYTRVYIYDMTHY 240
 QY 271 VPKKP 275
 DB 241 VPKKP 245

RESULT 13
 US-09-079-970A-5
 Sequence 5, Application US/09079970A
 Patent No. 6274366
 GENERAL INFORMATION:
 APPLICANT: Maflet, Mark A.
 APPLICANT: Miles, Andrew L.
 APPLICANT: Haak-Frendscho, Mary
 TITLE OF INVENTION: Enzymatically-Active Recombinant Human
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESS: Intellectual Property Department
 STREET: 8000 Excelstior Drive, Suite 401
 CITY: Madison
 STATE: WISCONSIN
 COUNTRY: U.S.A.
 ZIP: 53717-1914
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/079,970A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Leone, Joseph T.
 REGISTRATION NUMBER: 37,170
 REFERENCE/DOCKET NUMBER: 34506.073
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 831-2100
 TELEFAX: (608) 831-2100
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 249 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-079-970A-5

Query Match 84.9%; Score 1284; DB 3; Length 249;
 Best Local Similarity 92.7%; Pred. No. 1.5e-128;
 Matches 227; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 31 IYGGQEARSKMPQVSLRVRDRYMMHFCGSSLIHPQWVTLAAHCLGPDVYKDLATLRVQL 90
 DB 5 IYGGQEARSKMPQVSLRVRDRYMMHFCGSSLIHPQWVTLAAHCLGPDVYKDLATLRVQL 64
 QY 91 RQHLIYQDOLLPSRRIIVHPQFYIIOTGADIALLEBEPNVISRRVHTWMLPPASETFP 150
 DB 65 RQHLIYQDOLLPSRRIIVHPQFYIIOTGADIALLEBEPNVISRRVHTWMLPPASETFP 124
 QY 151 PGMPCWVTGMGDVNDDEPLPPFPPLKQVVPIMENHICDAKYHLGAYTGDVRIIRDML 210
 DB 125 PGMPCWVTGMGDVNDDEPLPPFPPLKQVVPIMENHICDAKYHLGAYTGDVRIIRDML 184
 QY 211 CAGNSQRDSCKDSGGGLVCKVNGTWTLAGVYVSWDEGCAQPNRPGIYTRVYIYDMTHY 270
 DB 185 CAGNTRRDS CGDGGGLVCKVNGTWTLAGVYVSWDEGCAQPNRPGIYTRVYIYDMTHY 244
 QY 271 VPKKP 275
 DB 245 VPKKP 249

RESULT 14
 US-09-601-318-4
 Sequence 4, Application US/09601318
 Patent No. 6613769
 GENERAL INFORMATION:
 APPLICANT: Max-Planck-Gesellschaft z. Fnd. d. Wissenssch. e.V
 APPLICANT: Byk Gulden Lomberg Chemische Fabrik GmbH
 APPLICANT: Bode, Wolfram
 APPLICANT: Moroder, Luis
 APPLICANT: Pereira, Pedro Jose Barbosa
 APPLICANT: Bergner, Andreas
 APPLICANT: Huber, Robert
 APPLICANT: Sommerhoff, Christian
 APPLICANT: Schaschke, No. 6613769bert
 APPLICANT: Br, Thomas
 APPLICANT: Martini, Thomas
 APPLICANT: Stadlwieser, Josef
 APPLICANT: Ulrich, Wolf-Rdiger
 APPLICANT: Dominik, Andreas
 APPLICANT: Thibaut, Ulrich
 APPLICANT: Bundschuh, Daniela
 APPLICANT: Beume, Rolf
 APPLICANT: Goebel, Karl-Josef
 TITLE OF INVENTION: Typlase-Inhibitoren
 FILE REFERENCE: 17674P WO-1
 CURRENT APPLICATION NUMBER: US/09/601,318
 CURRENT FILING DATE: 2001-01-22
 PRIOR APPLICATION NUMBER: DE19804761.4
 PRIOR FILING DATE: 1998-02-06
 PRIOR APPLICATION NUMBER: DE19851300.3
 PRIOR FILING DATE: 1998-11-06
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 244
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-601-318-4

Query Match 84.5%; Score 1277; DB 4; Length 244;
 Best Local Similarity 92.6%; Pred. No. 7.9e-128;
 Matches 226; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

```

QY 31 IYGGQEARPSKMPWQVSLRVYRDRYMMHFCGGSLIHPQWVLTAAHCGIPVDKDIATLRVOL 90
DB 1 IYGGQEARPSKMPWQVSLRVYRDRYMMHFCGGSLIHPQWVLTAAHCGIPVDKDIATLRVOL 60
QY 91 REQHLYYQDQDLBVSRIIVHPQFYIIQTAGADIALLEBEPVNISSRVHTVMLPPASSTFP 150
DB 61 REQHLYYQDQDLBVSRIIVHPQFYIIQTAGADIALLEBEPVNISSRVHTVMLPPASSTFP 120
QY 151 PGMPCWVTGMDVNDERLPPRPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIVRDDML 210
DB 121 PGMPCWVTGMDVNDERLPPRPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIVRDDML 180
QY 211 CAGNSQRDSCCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 270
DB 181 CAGNTRRDSCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 240
QY 271 VPKK 274
DB 241 VPKK 244

```

RESULT 15

```

US-09-601-318-5
; Sequence 5, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft z. Ftd. d. Wissenssch. e.V
; APPLICANT: Byk Gulden Lomberg Chemische Fabrik GmbH
; APPLICANT: Bode, Wolfram
; APPLICANT: Moroder, Luis
; APPLICANT: Bereira, Pedro Jose Barbosa
; APPLICANT: Bergner, Andreas
; APPLICANT: Huber, Robert
; APPLICANT: Sommerhoff, Christian
; APPLICANT: Schaschke, No. 6613769bert
; APPLICANT: Br. Thomas
; APPLICANT: Martin, Thomas
; APPLICANT: Stadlmeier, Josef
; APPLICANT: Ulrich, Wolf-Ridiger
; APPLICANT: Dominik, Andreas
; APPLICANT: Thibaut, Ulrich
; APPLICANT: Bundeschuh, Daniela
; APPLICANT: Beume, Rolf
; APPLICANT: Goebel, Karl-Josef
; TITLE OF INVENTION: Tryptase-Inhibitoren
; FILE REFERENCE: 17674P WO-1
; CURRENT APPLICATION NUMBER: US/09/601,318
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: DE19804761.4
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: DE19851300.3
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-601-318-5

```

```

Query Match 84.5%; Score 1277; DB 4; Length 244;
Best Local Similarity 92.6%; Pred. No. 7, 9e-128;
Matches 226; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
QY 31 IYGGQEARPSKMPWQVSLRVYRDRYMMHFCGGSLIHPQWVLTAAHCGIPVDKDIATLRVOL 90
DB 1 IYGGQEARPSKMPWQVSLRVYRDRYMMHFCGGSLIHPQWVLTAAHCGIPVDKDIATLRVOL 60
QY 91 REQHLYYQDQDLBVSRIIVHPQFYIIQTAGADIALLEBEPVNISSRVHTVMLPPASSTFP 150
DB 61 REQHLYYQDQDLBVSRIIVHPQFYIIQTAGADIALLEBEPVNISSRVHTVMLPPASSTFP 120
QY 151 PGMPCWVTGMDVNDERLPPRPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIVRDDML 210
DB 121 PGMPCWVTGMDVNDERLPPRPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIVRDDML 180
QY 211 CAGNSQRDSCCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 270
DB 181 CAGNTRRDSCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 240
QY 271 VPKK 274
DB 241 VPKK 244

```

```

DB 121 PGMPCWVTGMDVNDERLPPRPLKQVKVPIEMENHICDAKXHLGAYTGDDVRIVRDDML 180
QY 211 CAGNSQRDSCCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 270
DB 181 CAGNTRRDSCKGDSGGPILVCKVNGTWMLOAGVYVSWDEGCAQPRRPGIYTRVYYLDMIHNY 240
QY 271 VPKK 274
DB 241 VPKK 244

```

Search completed: August 27, 2005, 23:31:26
Job time: 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2005, 23:12:00 ; Search time 124 Seconds
(without alignments)
857.735 Million cell updates/sec

Title: US-09-598-982C-52

Perfect score: 1 MSLSLLALPVLASRVAAP.....ITTRVTVYLDWIMHHVPPKRP 275

Scoring table: BLOSSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq.16dec04:*
- 1: geneseqp1980s:**
 - 2: geneseqp1990s:**
 - 3: geneseqp2000s:**
 - 4: geneseqp2001s:**
 - 5: geneseqp2002s:**
 - 6: geneseqp2003as:**
 - 7: geneseqp2003bs:**
 - 8: geneseqp2004s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1512	100.0	275	2 AAW64237	AAW64237 Human mas
2	1512	100.0	275	2 AAW63173	AAW63173 Human mas
3	1512	100.0	275	7 ADKS2388	ADKS2388 Hematolog
4	1512	100.0	275	8 ADP12448	ADP12448 Protein e
5	1512	100.0	275	8 ADP56070	ADP56070 Human PRO
6	1505	99.5	275	7 ADSE62892	ADSE62892 Human PRO
7	1430	94.6	275	7 ADSE56171	ADSE56171 Human PRO
8	1411	93.3	275	7 ADJ94864	ADJ94864 Novel NOV
9	1411	93.3	275	8 ADQ19871	ADQ19871 Human sof
10	1405	92.9	275	7 ADJ94866	ADJ94866 Novel NOV
11	1405	92.9	275	7 ADKS2590	ADKS2590 Hematolog
12	1402	92.7	273	2 AAW64238	AAW64238 Human mas
13	1402	92.7	273	2 AAW63174	AAW63174 Human mas
14	1401	92.7	691	5 AAE14348	AAE14348 Human PRO
15	1400	92.6	274	2 AAW64240	AAW64240 Human mas
16	1400	92.6	274	2 AAW63175	AAW63175 Human mas
17	1394	92.2	267	2 AAW64241	AAW64241 Human mas
18	1394	92.2	267	2 AAW63176	AAW63176 Human mas
19	1394	92.2	267	5 AAUS84360	AAUS84360 Protein T
20	1382	91.4	275	7 ADJ94862	ADJ94862 Novel NOV
21	1360	89.9	264	7 ADG42743	ADG42743 Human bet
22	1360	89.9	264	8 ADM76650	ADM76650 Human NOV
23	1349	89.2	266	7 ADG42723	ADG42723 Human mas
24	1349	89.2	266	8 ADM76630	ADM76630 Human NOV
25	1346.5	89.1	333	8 ADJ72001	ADJ72001 Human PWM

ALIGNMENTS

RESULT 1	ID	AAW64237 standard; protein, 275 AA.	ADG42731 Human bet
26	1324	87.6	256 7
27	1324	87.6	256 8
28	1323	87.5	264 7
29	1316	87.0	254 7
30	1313	86.8	254 7
31	1313	86.8	254 8
32	1307	85.4	256 7
33	1298.5	85.9	275 7
34	1290	85.3	245 5
35	1290	85.3	249 5
36	1287	85.1	245 5
37	1287	85.1	245 5
38	1287	85.1	249 5
39	1287	85.1	249 5
40	1284	84.9	245 5
41	1284	84.9	249 3
42	1284	84.9	249 5
43	1282	84.8	245 5
44	1282	84.8	249 5
45	1281	84.7	245 5

AAW64237
ID AAW64237 standard; protein, 275 AA.
AC AAW64237;
XX
XX
DT 24-NOV-1998 (first entry)
XX
XX
DE Human mast cell tryptase alpha.
KW Mast cell tryptase alpha; human; MCP-7; mast cell protease 7; blood clot; anticogulant; myocardial infarction; reocclusion; thromboembolism; cerebral embolism; thrombosis; therapy.
XX
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..20
FT Peptide /label= Sig_peptide
FT Peptide 21..30
FT Protein /label= Pro_peptide
FT Protein /label= Mat_protein

W09824886-A1.
11-JUN-1998.
25-NOV-1997; 97WO-US021620.
04-DEC-1996; 96US-0032354P.
(BGHM) BRIGHAM & WOMENS HOSPITAL.
Stevens RL;
WPI, 1998-333308/29.
N-PSDB; AAV44328.

New compositions containing tryptase-7, e.g. mouse mast cell protease-7 are used to treat clot formation in e.g. myocardial infarction, reocclusion following angioplasty or pulmonary thrombo-embolism.

Disclosure: Page 63-64; 92pp; English.

This is the deduced amino acid sequence of human mast cell tryptase alpha (see also AAV44328). The invention provides compositions comprising an isolated tryptase-7 that may include chimeric proteins that contain (a) a

CC human tryptase for all but the active site region and (b) the substrate-
 CC binding pocket of mouse tryptase-7 or its homologues (see AAW64233-39); a
 CC method for treating a blood clot by administering a nucleic acid molecule
 CC that codes for a tryptase-7, or an expression product, to decrease
 CC fibrinogen activity; a nucleic acid encoding a serine protease (SP); and
 CC a method of producing a mature SP by expressing the inactive zymogen in a
 CC host cell, and cleaving the enterokinase susceptibility domain. The
 CC tryptase-7 polypeptides can be used to treat disorders mediated by
 CC undesirable thrombus clot formation such as myocardial infarction and
 CC reconstruction following angioplasty of blood clots associated with
 CC pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal
 CC vein and peripheral arterial thrombosis. They are also useful for all
 CC surgical procedures that require decreased blood clots

SO Sequence 275 AA;
 Query Match 100.0%; Score 1512; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLLILALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPWQVSLRVDRYMMHFCG 60
 Db 1 MSLLILALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPWQVSLRVDRYMMHFCG 60
 QY 61 GSIHPQWVLTAAHCLGPPVDKDLATLRVQLREQHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 Db 61 GSIHPQWVLTAAHCLGPPVDKDLATLRVQLREQHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASSTFPFGMPQWVTGMDVNDDEPLPEPFLKQYKV 180
 Db 121 DIALLELEBPVNISSRVHTVMLPPASSTFPFGMPQWVTGMDVNDDEPLPEPFLKQYKV 180
 QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCGKDSGSPVYCKVNGTMLQAG 240
 Db 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCGKDSGSPVYCKVNGTMLQAG 240
 QY 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHHVYPKP 275
 Db 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHHVYPKP 275

RESULT 2
 AAW63173
 ID AAW63173 standard; protein; 275 AA.
 AC AAW63173;
 XX
 DT 27-OCT-1998 (first entry)

DE Human mast cell tryptase alpha polypeptide.
 DE
 XX Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
 KM trypsinase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
 KM antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
 KM hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
 KM inflammatory skin condition; human; mast cell tryptase alpha.

OS Homo sapiens.
 XX
 PN W09833812-A1.
 XX
 PD 06-AUG-1998.
 XX
 PF 30-JAN-1998; 98MO-US001865.
 XX
 PR 05-FEB-1997; 97US-0037090P.
 XX

PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Stevens RL, Huang C;
 XX
 DR WPI; 1998-437390/37.
 DR N-PSDB; AAV42710.
 DR

XX Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated
 PR inflammatory disorders e.g. asthma.
 PT
 XX Discloure; Page 45-46; 69pp; English.
 PS
 XX This reorganizes the human mast cell tryptase alpha protein which is a
 CC homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention
 CC provides sequences shown in AAW63160 to AAW63169 that are inhibitors of
 CC mMCP-6. These peptides which are tryptase-6 complex inhibitors, can be
 CC used for treating a mast cell-mediated inflammatory disorder. The
 CC inhibitors can be used to treat inflammatory disorders including asthma,
 CC allergic rhinitis, urticaria and antioedema, eczematous dermatitis
 CC (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic
 CC ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory
 CC skin conditions

SO Sequence 275 AA;
 Query Match 100.0%; Score 1512; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLLILALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPWQVSLRVDRYMMHFCG 60
 Db 1 MSLLILALPVLASRAVAAPAVQALQOAGIVGGGEAPRSKMPWQVSLRVDRYMMHFCG 60
 QY 61 GSIHPQWVLTAAHCLGPPVDKDLATLRVQLREQHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 Db 61 GSIHPQWVLTAAHCLGPPVDKDLATLRVQLREQHLYYQDQLLPSRRIIVHPQFYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASSTFPFGMPQWVTGMDVNDDEPLPEPFLKQYKV 180
 Db 121 DIALLELEBPVNISSRVHTVMLPPASSTFPFGMPQWVTGMDVNDDEPLPEPFLKQYKV 180
 QY 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCGKDSGSPVYCKVNGTMLQAG 240
 Db 181 PIMENHICDAKXHLGAYTGDDVRIIRDMLCAGNSQRDSCGKDSGSPVYCKVNGTMLQAG 240
 QY 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHHVYPKP 275
 Db 241 VVSMDEGCAQPNRPGIYTRVTVYLLDWHIHHVYPKP 275

RESULT 3
 ADK52588
 ID ADK52588 standard; protein; 275 AA.
 AC ADK52588;
 XX
 DT 06-MAY-2004 (first entry)

DE Hematological disorder associated Gene ID 1847 encoded protein.
 DE
 XX
 XX cytostatic; antianemic; antistickling; vitruclide; hemostatic; nephrotropic;
 KM cytoslastic; thrombolytic; antiparasitic; gene therapy;
 KM hematologic disorder; cancer; Sickle Cell Anemia;
 KM Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
 KM Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
 KM transfusion reaction; Erythroblastosis; mechanical trauma;
 KM micro-angiopathic hemolytic anemia; parasite infection.

OS Homo sapiens.
 XX
 PN W02003065871-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 28-JAN-2003; 2003WO-US002484.
 XX
 PR 04-FEB-2002; 2002US-0354333P.
 PR 28-FEB-2002; 2002US-0360258P.
 PR 15-MAR-2002; 2002US-0364476P.
 PR

26-APR-2002; 2002US-0375626P.
 PR 06-JUN-2002; 2002US-0386494P.
 PR 24-JUN-2002; 2002US-0390965P.
 PR 28-JUN-2002; 2002US-0392480P.
 PR 03-JUL-2002; 2002US-0394128P.
 PR 31-JUL-2002; 2002US-0399783P.
 PR 13-AUG-2002; 2002US-0403221P.
 PR 30-AUG-2002; 2002US-0407045P.
 PR 25-NOV-2002; 2002US-0429048P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Carcoll JM, Healy A, Welch NS, Kelly LM;
 PI WPI: 2003-731464/69.
 DR N-PSDB; ADKS2587.
 XX
 XX Identifying a compound capable of treating a hematologic disorder (e.g.
 PT anemia or leukemia) comprises assaying the ability of the compound to
 PR modulate the expression or activity of e.g. 131,148, 199 or 12303
 XX polypeptide or nucleic acid.
 PS Disclosure; SEQ ID NO 46; 232pp; English.
 XX
 CC The invention relates to a method of identifying a compound capable of
 CC treating a hematologic disorder comprising assaying the ability of the
 CC compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
 CC 194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
 CC 13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
 CC acid expression or polypeptide activity, thus, identifying a compound
 CC capable of treating a hematologic disorder. The methods are useful in
 CC diagnosing, preventing and treating hematological disorders, such as
 CC cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
 CC Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, Thalassemia,
 CC associated with an increased risk of Thrombosis, Herpes, Thalassemia,
 CC antibody-mediated disorders such as transfusion reactions and
 CC erythroblastosis, mechanical trauma to red blood cells such as micro-
 CC anglopathic hemolytic anemias, infections by parasites or chemical
 CC injuries. The methods may also be used for identifying compounds that
 CC modulate hematological disorders. This sequence corresponds to the
 CC protein encoded by one of the genes modulated by the compounds.
 XX
 XX Sequence 275 AA:
 SQ
 Query Match 100.0%; Score 1512; DB 7; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ADP12448;
 AC
 XX 12-AUG-2004 (first entry)
 DT
 XX Protein encoded by mRNA of the invention #58.
 DE
 XX transplant rejection; immune system; rheumatoid arthritis; lupus;
 XX inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
 KW
 XX Homo sapiens.
 OS
 XX
 XX WO2004042346-A2.
 XX
 XX 21-MAY-2004.
 XX
 XX 24-APR-2003; 2003WO-US012946.
 XX
 XX 24-APR-2002; 2002US-00131831.
 XX
 XX 20-DEC-2002; 2002US-00325899.
 XX
 XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
 PA
 XX
 XX Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
 PI Rosenbery S;
 XX
 XX WPI: 2004-400724/37.
 DR
 XX
 XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
 XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
 XX rejection, in an individual, comprises detecting the expression level of
 XX the genes.
 PS Claim 65; SEQ ID NO 2457; 1762pp; English.
 XX
 CC The present invention relates to diagnosing or monitoring transplant
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
 CC comprises detecting the expression level of one or more genes. The
 CC methods, system and kits are useful in diagnosing or monitoring
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
 CC islet, lung, bone marrow or stem cell transplant rejection,
 CC xenotransplant rejection or mechanical organ replacement rejection, in an
 CC individual. The method is also useful in assessing the immune status of
 CC an individual. The methods are also useful in diagnosing and monitoring
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
 CC viral, bacterial or fungal infection. The present sequence represents a
 CC protein that is encoded by the mRNA of the invention.
 XX
 XX Sequence 275 AA:
 SQ
 Query Match 100.0%; Score 1512; DB 8; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 ADP56070 standard; protein; 275 AA.
 ADP56070;
 18-NOV-2004 (first entry)
 Human PRO protein sequence SEQ ID NO:2046.
 human; PRO; immune related disease; inflammatory immune response;
 immune response stimulation; anti-allergic; anti-nausea; antiarthritic;
 antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
 antirheumatic; antihydroid; CNS; dermatological; gastrointestinal;
 haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 virucide; gene therapy.
 Homo sapiens.
 WO2004039956-A2.
 13-MAY-2004.
 28-OCT-2003; 2003WO-US034381.
 29-OCT-2002; 2002US-0422472P.
 (GENTEC) GENENTECH INC.
 Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM,
 Wood WI, Wu TD;
 MPI: 2004-376182/35.
 N-FSDB; ADP56069.
 New PRO polynucleotides and polypeptides, useful in useful in diagnosing
 and treating an immune related disease, e.g. systemic lupus
 erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 stimulating an immune response.
 Claim 1; SEQ ID NO 2046; 30099pp; English.

Sequence 275 AA:
 Query Match 100.0%; Score 1512; DB 8; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSLILLALPVLASRAVAAPAVQALQOAGIVGCGEAPRSKWPQVSLRVRDRYMMHFCG 60
 DB 1 MSLILLALPVLASRAVAAPAVQALQOAGIVGCGEAPRSKWPQVSLRVRDRYMMHFCG 60
 QY 61 GSLIHPQWLTFAHCLGPPVVDLTLRVQLREQHLIYDQDLIPVSRITIVHPQYIIQTGA 120
 DB 61 GSLIHPQWLTFAHCLGPPVVDLTLRVQLREQHLIYDQDLIPVSRITIVHPQYIIQTGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCVTWGWDVNDDEPLPFPPLKQVKV 180
 DB 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPCVTWGWDVNDDEPLPFPPLKQVKV 180
 QY 181 PIMENHICDAKYLHLGAYGGDVRITRDMLCAGNSQRDSCKDGSGGPLVCKVNGTWLQAG 240
 DB 181 PIMENHICDAKYLHLGAYGGDVRITRDMLCAGNSQRDSCKDGSGGPLVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPNRPGIYTRVYVYLLMHHVYVPKKP 275
 DB 241 VVSWDEGCAQPNRPGIYTRVYVYLLMHHVYVPKKP 275
 RESULT 6
 ADE62892
 ID ADE62892 strandard; protein; 275 AA.
 AC ADE62892;
 DT 29-JAN-2004 (first entry)
 DE Human Protein P15157, SEQ ID NO 8826.
 XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KM spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PR
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M;
 DR MPI: 2003-268312/26.
 DR GENBANK; P15157.
 XX
 PS Claim 1; Page: 1017pp; English.
 The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 derivative or allelic variation of the nucleic acid sequence. Also
 claimed are a vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 which is differentially regulated in an animal subjected to pain and a
 kit to perform the method, an array, a method for identifying an agent

that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 99.5%; Score 1505; DB 7; Length 275;
 Best Local Similarity 99.3%; Pred. No. 5,5e-133;
 Matches 273; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 MLSLILALPVLASRAVAAPAPVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60
 1 MLNLIILALPVLASRAVAAPAPVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275

RESULT 7
 ADE56171
 ID ADE56171 standard; protein; 275 AA.
 AC ADE56171;
 DT 29-JAN-2004 (first entry)
 DE Human Protein XP_018104, SEQ ID NO 2020.
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 OS Homo sapiens.
 PN WO2003016475-A2.
 PD 27-FEB-2003.
 PF 14-AUG-2002; 2002MO-US025765.
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.

XX (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan W,
 DR WPI: 2003-268312/26.
 DR GENBANK; XP_018104.
 PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 PS Claim 1; Page: 1017p; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 94.6%; Score 1430; DB 7; Length 275;
 Best Local Similarity 93.5%; Pred. No. 4.9e-124;
 Matches 257; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

1 MSLIILALPVLASRAVAAPAPVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60
 1 MLNLIILALPVLASRAVAAPAPVQALQOAGIVGGGEAPRSKMPQVSLRVDRYMMHFCG 60
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 61 GSLLHPQWVLTAAHCLGPDVVDLALRVQLREQHLYYODQLLPSRIIVHQPFIYIOTGA 120
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 121 DIALLEEPEPVNISRVTVMTPPASETFPPGMPGCVTGMGVDVNDDEPLPPFPPLKQVKV 180
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 181 PIMENHICDAKYHLAGYTGDDVRIIRDDMLCAGNSORDSCGDSGGPPLVCKVNGTWTLOAG 240
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275
 241 VVSWDEGCAQPNRPGIYTRVYIYLDWIHHYVPKKP 275

RESULT 8
 ADJ94864
 ID ADJ94864 standard; protein; 275 AA.

```

AC ADJ94864;
XX
XX 06-MAY-2004 (first entry)
XX Novel NOVX protein sequence #46.
XX
XX antidiabetic; anorectic; cardiact; hypotensive; antiarteriosclerotic;
XX anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
XX neuroprotective; antiparkinsonian; osteoconvulsant; osteopathic;
XX antiarthritic; antiinflammatory; dermatological; antiasthmatica;
XX antiinflammatory; gene therapy; metabolic disorder; diabetes; obesity;
XX infectious disease; anorexia; cancer; cardiovascular disease;
XX hypertension; atherosclerosis; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
XX osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
XX asthma; dyslipidemia; neurogenesis; cell differentiation;
XX cell proliferation; hematopoiesis; wound healing; angiogenesis;
XX chromosome mapping; tissue typing; pharmacogenomic.
XX
XX Homo sapiens.
XX
XX WO2003040325-A2.
XX
XX 15-MAY-2003.
XX
XX 05-NOV-2002; 2002W0-US035464.
XX
XX 05-NOV-2001; 2001US-0338626P.
XX
XX 06-NOV-2001; 2001US-033072P.
XX
XX 09-NOV-2001; 2001US-0348283P.
XX
XX 15-NOV-2001; 2001US-0335610P.
XX
XX 16-NOV-2001; 2001US-0338543P.
XX
XX 20-NOV-2001; 2001US-0331631P.
XX
XX 20-NOV-2001; 2001US-0331641P.
XX
XX 21-NOV-2001; 2001US-0332152P.
XX
XX 27-NOV-2001; 2001US-0333461P.
XX
XX 28-NOV-2001; 2001US-0333912P.
XX
XX 28-NOV-2001; 2001US-0334027P.
XX
XX 29-NOV-2001; 2001US-0334300P.
XX
XX 30-NOV-2001; 2001US-0334421P.
XX
XX 31-NOV-2001; 2001US-0334526P.
XX
XX 04-DEC-2001; 2001US-0336576P.
XX
XX 04-DEC-2001; 2001US-0336664P.
XX
XX 07-DEC-2001; 2001US-0338314P.
XX
XX 07-DEC-2001; 2001US-0338390P.
XX
XX 10-DEC-2001; 2001US-0339006P.
XX
XX 11-DEC-2001; 2001US-0339286P.
XX
XX 01-FEB-2002; 2002US-0353280P.
XX
XX 01-FEB-2002; 2002US-0353288P.
XX
XX 04-FEB-2002; 2002US-0354392P.
XX
XX 04-FEB-2002; 2002US-0354393P.
XX
XX 04-FEB-2002; 2002US-0354409P.
XX
XX 27-FEB-2002; 2002US-0359944P.
XX
XX 27-FEB-2002; 2002US-0360148P.
XX
XX 05-MAR-2002; 2002US-0361790P.
XX
XX 05-MAR-2002; 2002US-0361833P.
XX
XX 05-MAR-2002; 2002US-0361925P.
XX
XX 05-MAR-2002; 2002US-0362230P.
XX
XX 05-MAR-2002; 2002US-0362625P.
XX
XX 13-MAR-2002; 2002US-0364000P.
XX
XX 13-MAR-2002; 2002US-0364181P.
XX
XX 13-MAR-2002; 2002US-0364182P.
XX
XX 13-MAR-2002; 2002US-0364197P.
XX
XX 17-MAY-2002; 2002US-0381621P.
XX
XX 28-MAY-2002; 2002US-0383675P.
XX
XX 17-THI-2002; 2002US-0396703P.
XX
XX 06-AUG-2002; 2002US-0401552P.
XX
XX 07-AUG-2002; 2002US-0401594P.
XX
XX 15-AUG-2002; 2002US-0401787P.
XX
XX 20-AUG-2002; 2002US-0403619P.
XX
XX 20-AUG-2002; 2002US-0404821P.

```

```

PR 23-AUG-2002; 2002US-0405368P.
PR 23-AUG-2002; 2002US-0405402P.
PR 23-AUG-2002; 2002US-0405496P.
PR 23-AUG-2002; 2002US-0405631P.
PR 26-AUG-2002; 2002US-0406125P.
PR 04-NOV-2002; 2002US-00287226.
XX
XX (CURA-) CURAGEN CORP.
XX
XX PA
XX PI Agee ML, Alsobrook JP, Bergna C, Boldog FL, Burgess CE, Chant JS;
XX PI Chandhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K,
XX PI Gangolli EA, Gorman L, Garlach VL, Ji W, Kekuda R, Khrantsov NY,
XX PI Li L, Malyanar UM, Macdougall JR, Mezes PS, Miller CE, Millet I,
XX PI Oei GE, Ort T, Padigaru M, Patutarajan M, Raselli L, Rieger DK,
XX PI Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ,
XX PI Vernet CM, Zernusen BD, Zhong M;
XX
XX DR MPI: 2003-441551/41.
XX DR N-PSDB; ADJ94863.
XX
XX PT New isolated NOVX polypeptides and polynucleotides, useful for
XX PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
XX PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
XX PT asthma, or infections.
XX
XX PS Claim 1; SEQ ID NO 92; 800pp; English.
XX
XX
XX The invention relates to novel isolated polypeptides, mature forms of
XX these, or a sequence that is at least 95 % identical to, or having one or
XX more conservative amino acid substitutions in the polypeptides. The
XX polypeptides, nucleic acid molecules and antibodies are useful in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease, preferably a NOVX-associated disorder. The nucleic acid
XX molecules, polypeptides and antibodies are useful for treating,
XX preventing or diagnosing diseases such as metabolic disorders, diabetes,
XX obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
XX protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
XX atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
XX Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
XX hematopoietic disorders, inflammatory skin disorders, asthma, and various
XX dyslipidemias. The nucleic acids and polypeptides may also be used as
XX targets for the identification of small molecules that modulate or
XX inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
XX hemtopoiesis, wound healing and angiogenesis, in gene therapy, in
XX generation of antibodies that bind immunospecifically to NOVX substances
XX for use in therapeutic or diagnostic methods. The nucleic acids are
XX further used as hybridization probes, in chromosome mapping, tissue
XX typing, preventive medication, and pharmacogenomics. This sequence
XX corresponds to one of the NOVX polypeptides of the invention.
XX
XX SQ Sequence 275 AA;
XX
XX Query Match 93.3%; Score 1411; DB 7; Length 275;
XX Best local Similarity 92.4%; Pred. No. 2,9e-122;
XX Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;
XX
XX 1 MSLDLLALPVLASPAVAAAPVQALQAGVGGGEAPRSKRWQVSLRVDRYMMHFCG 60
XX | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 MNLMLLALPVLASPAVAAAPVQALQAGVGGGEAPRSKRWQVSLRVDRYMMHFCG 60
XX
XX QY 61 GSLLHPQWLITAHNCLGPRVVDLALTRVQLREKOHLYYDQDLIIVRSRIITVHPQFYITIGGA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GSLLHPQWLITAHNCLGPRVVDLALTRVQLREKOHLYYDQDLIIVRSRIITVHPQFYITIGGA 120
XX
XX Db DIALLELEEPVNISSRVTVMILPASETFRPPMPCWVTGWDVNDDELPFPPFLKQVKV 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 DIALLELEEPVNVSSHVHTVTLPASFTFRPPMPCWVTGWDVNDDELPFPPFLKQVKV 180
XX
XX QY 181 PIMENHICDAKYHLGAYTGDVRIITRDMVLKAGNSQKDSCKKDGSGPILVCKYKNTWLQAG 240
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 PIMENHICDAKYHLGAYTGDVRIIVRDMVLKAGNTRKRDSCQDGSQGPLVCKYKNTWLQAG 240
XX
XX Db 181 PIMENHICDAKYHLGAYTGDVRIIVRDMVLKAGNTRKRDSCQDGSQGPLVCKYKNTWLQAG 240
XX
XX QY 241 VVSNDEGCAQPRRPGITRYRVTVYUULDIHNVYKPK 275

```

DB 241 VVSWGEGCAQPNRPGIYTRVYIYLDWTHHYVPEKRP 275

RESULT 9
ADQ19871
ID ADQ19871 standard; protein; 275 AA.

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2690.

KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
KW Homo sapiens.

OS WO2004048938-A2.

PN 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnick A;

XX WPI; 2004-441208/41.

PT Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

PS Example 2; SEQ ID NO 2690; 210pp; English.

CC The invention relates to a novel method for detecting soft tissue sarcoma and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

SO Sequence 275 AA;

Query Match 93.3%; Score 1411; DB 8; Length 275;

Best Local Similarity 92.4%; Pred. No. 2,9e-122;

Matches 254; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 1 MSLILLALPVLASRYAAPVAPVQALQAGIYGGGBAPRSKMPWQVSLAVRDRYMMHFFG 60
DB 1 MNLILLALPVLASRYAAPVAPVQALQAGIYGGGBAPRSKMPWQVSLVHGRPYMMHFFG 60
OY 61 GSIILHQPWVLTAAHCLGPDVKIATLRVQIREQHLTYQDQLLPSKRIIVHPQFYIIQTGA 120
DB 61 GSIILHQPWVLTAAHCLGPDVKIATLRVQIREQHLTYQDQLLPSKRIIVHPQFYIIQTGA 120
OY 121 DIALLELEBPPVNISSRYVTVMLPPASSETPPPCMPVNTGKGDVNDDEPLPPPPPLKQYKV 180
DB 121 DIALLELEBPPVNISSRYVTVMLPPASSETPPPCMPVNTGKGDVNDDEPLPPPPPLKQYKV 180
OY 181 PIMENHICDAKYHLGAVYTGDDVRIIRDDMLCAGNSQRDSCKGGSGPLVCKVNGTWLQAG 240
DB 181 PIMENHICDAKYHLGAVYTGDDVRIIRDDMLCAGNSQRDSCKGGSGPLVCKVNGTWLQAG 240

DB 181 PIMENHICDAKYHLGAVYTGDDVRIIRDDMLCAGNSQRDSCGDSGGLVCKVNGTWLQAG 240
OY 241 VVSWDEGCAQPNRPGIYTRVYIYLDWTHHYVPEKRP 275
DB 241 VVSWGEGCAQPNRPGIYTRVYIYLDWTHHYVPEKRP 275

RESULT 10
ADJ94866
ID ADJ94866 standard; protein; 275 AA.

AC ADJ94866;

DI 06-MAY-2004 (first entry)

DE Novel NOVX protein sequence #47.

KW antidiabetic; anorectic; cardiatic; hypotensive; antiarteriosclerotic;
KW anorectic; virocidic; antibacterial; fungicide; protozoacide; nootropic;
KW neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;
KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
KW antilipemic; gene therapy; metabolic disorder; diabetes; obesity;
KW infectious disease; anorexia; cancer; cardiovascular disease;
KW hypertension; atherosclerosis; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
KW osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
KW asthma; dyslipidemia; neurogenesis; cell differentiation;
KW cell proliferation; hematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; pharmacogenomic.

OS Homo sapiens.

PN WO2003040325-A2.

PD 15-MAY-2003.

PF 05-NOV-2002; 2002WO-US035464.

XX 05-NOV-2001; 2001US-0338626P.
XX 06-NOV-2001; 2001US-0333072P.
XX 09-NOV-2001; 2001US-0348283P.
XX 15-NOV-2001; 2001US-0335610P.
XX 16-NOV-2001; 2001US-0338543P.
XX 20-NOV-2001; 2001US-0331630P.
XX 20-NOV-2001; 2001US-0331641P.
XX 21-NOV-2001; 2001US-0332152P.
XX 27-NOV-2001; 2001US-0334461P.
XX 28-NOV-2001; 2001US-0334027P.
XX 28-NOV-2001; 2001US-0334027P.
XX 29-NOV-2001; 2001US-0334300P.
XX 30-NOV-2001; 2001US-0334421P.
XX 30-NOV-2001; 2001US-0334526P.
XX 04-DEC-2001; 2001US-0336576P.
XX 04-DEC-2001; 2001US-0336664P.
XX 07-DEC-2001; 2001US-0338314P.
XX 07-DEC-2001; 2001US-0338390P.
XX 10-DEC-2001; 2001US-0339006P.
XX 10-DEC-2001; 2001US-0339008P.
XX 11-DEC-2001; 2001US-0339286P.
XX 01-FEB-2002; 2002US-0353280P.
XX 01-FEB-2002; 2002US-0353288P.
XX 04-FEB-2002; 2002US-0354322P.
XX 04-FEB-2002; 2002US-0354393P.
XX 04-FEB-2002; 2002US-0354409P.
XX 27-FEB-2002; 2002US-0359944P.
XX 27-FEB-2002; 2002US-0360148P.
XX 05-MAR-2002; 2002US-0361790P.
XX 05-MAR-2002; 2002US-0361833P.
XX 05-MAR-2002; 2002US-0361925P.
XX 05-MAR-2002; 2002US-0362230P.
XX 13-MAR-2002; 2002US-0364625P.
XX 13-MAR-2002; 2002US-0364000P.
XX 13-MAR-2002; 2002US-0364181P.

PR 13-MAR-2002; 2002US-0364182P.
 PR 13-MAR-2002; 2002US-0364197P.
 PR 13-MAR-2002; 2002US-0364227P.
 PR 17-MAY-2002; 2002US-0381621P.
 PR 28-MAY-2002; 2002US-0383675P.
 PR 17-JUL-2002; 2002US-0396703P.
 PR 06-AUG-2002; 2002US-0401552P.
 PR 07-AUG-2002; 2002US-0401594P.
 PR 15-AUG-2002; 2002US-0401787P.
 PR 20-AUG-2002; 2002US-0403619P.
 PR 23-AUG-2002; 2002US-0405368P.
 PR 23-AUG-2002; 2002US-0405402P.
 PR 23-AUG-2002; 2002US-0405496P.
 PR 23-AUG-2002; 2002US-0405631P.
 PR 26-AUG-2002; 2002US-0406125P.
 PR 04-NOV-2002; 2002US-00287226.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Agee M, Alsbrook JF, Berghe C, Boldog FL, Burgess CE, Chant JS;
 PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Elleman K;
 PI Gangoli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov AV;
 PI Li L, Malyaner UM, MacDougall JR, Mezes PS, Miller CE, Millet I;
 PI Ooi CE, Ort T, Padigar M, Patrajan M, Rastelli L, Rieger DK;
 PI Rohnberg ME, Shenoy SG, Spaderna SK, Spletter KA, Taupier RJ;
 PI Vernet CAM, Zernsen BD, Zhong M;
 XX
 DR WPI; 2003-441551/41.
 DR N-PSDB; ADJ94865.
 XX
 PT New isolated NOXV polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOXV-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 94; 800pp; English.
 XX
 CC The invention relates to novel isolated polypeptides, mature forms of
 CC these, or a sequence that is at least 95 % identical to, or having one or
 CC more conservative amino acid substitutions in the polypeptides. The
 CC polypeptides, nucleic acid molecules and antibodies are useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease, preferably a NOXV-associated disorder. The nucleic acid
 CC molecules, polypeptides and antibodies are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
 CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
 CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
 CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
 CC dyslipidemias. The nucleic acids and polypeptides may also be used as
 CC targets for the identification of small molecules that modulate or
 CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
 CC hemopoiesis, wound healing and angiogenesis, in gene therapy, in
 CC generation of antibodies that bind immunospecifically to NOXV substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridization probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. This sequence
 CC corresponds to one of the NOXV polypeptides of the invention.
 XX
 SO Sequence 275 AA;
 Query Match 92.9%; Score 1405; DB 7; Length 275;
 Best Local Similarity 92.0%; Pred. No. 1e-121;
 Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Db 61 GSLIHPQWVLTAAHCGPDKVQDLATLRVQLRQHLVYDQQLPVSRVITVHPQFYIAQIGA 120
 QY 121 DIALLELEBPVNISSRVHTVMLPPASETFPPGMPQCVTGWGDVNDDEPLPPFPKQVRY 180
 Db 121 DIALLELEBPVAVSSHHVHTVTLPPASETFPPGMPQCVTGWGDVNDDEPLPPFPKQVRY 180
 QY 181 PIMENHICDAKTYHLGAYGGDVRITRDMLCAGNSQRDSCKADSGGPIVCKVNGTWLQAG 240
 Db 181 PIMENHICDAKTYHLGAYGGDVRIVRDMLCAGNTRRSDSCGDSGGPIVCKVNGTWLQAG 240
 QY 241 VVSWDEGCAQPPRPGITRTVWYVYLLDMLHHVYVKKP 275
 Db 241 VVSWDEGCAQPPRPGITRTVWYVYLLDMLHHVYVKKP 275
 RESULT 11
 ADKS2590
 ID ADKS2590 standard; protein; 275 AA.
 AC ADKS2590;
 AC 06-MAY-2004 (first entry)
 DT
 DE Hematological disorder associated Gene ID 1849 encoded protein.
 XX
 KW cytostatic; anti-anemic; antisticking; virucide; hemostatic; nephrotropic;
 KW cytosolic; thrombolytic; antiparasitic; gene therapy;
 KW hematologic disorder; cancer; Sickle Cell Anemia;
 KW Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
 KW Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
 KW transfusion reaction; Erythroblastosis; mechanical trauma;
 KW micro-angiopathic hemolytic anemia; parasite infection.
 XX
 OS Homo sapiens.
 XX
 PN MO2003065871-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 28-JAN-2003; 2003WO-US002484.
 XX
 PP 04-FEB-2002; 2002US-0354333P.
 PR 28-FEB-2002; 2002US-0360258P.
 PR 15-MAR-2002; 2002US-0364476P.
 PR 26-APR-2002; 2002US-0375626P.
 PR 06-JUN-2002; 2002US-0386494P.
 PR 24-JUN-2002; 2002US-0390965P.
 PR 28-JUN-2002; 2002US-0392480P.
 PR 03-JUL-2002; 2002US-0394128P.
 PR 31-JUL-2002; 2002US-0399783P.
 PR 13-AUG-2002; 2002US-0403221P.
 PR 30-AUG-2002; 2002US-0407045P.
 PR 25-NOV-2002; 2002US-0429048P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Carroll JM, Healy A, Welch NS, Kelly LM;
 DR N-PSDB; ADKS2589.
 DR
 DR
 XX
 PT Identifying a compound capable of treating a hematologic disorder (e.g.
 PT anemia or leukemia) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 131,148, 199 or 12303
 PT polypeptide or nucleic acid.
 XX
 PS Disclosure; SEQ ID NO 48; 232pp; English.
 XX
 CC The invention relates to a method of identifying a compound capable of
 CC treating a hematologic disorder comprising assaying the ability of the
 CC compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
 CC 194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
 CC 13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic

CC acid expression or polypeptide activity, thus, identifying a compound
CC capable of treating a hematologic disorder. The methods are useful in
CC diagnosing, preventing and treating hematological disorders, such as
CC cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophiles, disorders
CC associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC antibody-mediated disorders such as transfusion reactions and
CC Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC angio-pathic hemolytic anemias, infections by parasites or chemical
CC injuries. The methods may also be used for identifying compounds that
CC modulate hematological disorders. This sequence corresponds to the
CC protein encoded by one of the genes modulated by the compounds.
XX Sequence 275 AA;

Query Match 92.9%; Score 1405; DB 7; Length 275;
Best Local Similarity 92.0%; Pred. No. 1e-121;
Matches 253; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSLTLLALPVLASRAVAAPVQALQOAGIVGGQEARPKMPOVSLRVRDRYMMHFCG 60
DB 1 MNLTLALPVLASRAVAAPVQALQOAGIVGGQEARPKMPOVSLRVHGPMWHFCG 60
QY 61 GSLIHPQWVLTAAHCIGRPDVKDLATLRVQLREONLHYODQLPVSRITVHPQFYIITQGA 120
DB 61 GSLIHPQWVLTAAHCIGRPDVKDLALRVQLREONLHYODQLPVSRITVHPQFYIITQGA 120
QY 121 DIALLEBEPVNISSRVHTVMLPPASERPPGMPGVDNDEPLRPPLKQVPI 180
DB 121 DIALLEBEPVNISSRVHTVMLPPASERPPGMPGVDNDEPLRPPLKQVPI 180
QY 181 PLMENHICDAKYNHLAGAVTGDDVRIIRDVMLCAGNSQRDSCGSPGVLCVKVGTWLQAGV 240
DB 181 PLMENHICDAKYNHLAGAVTGDDVRIIRDVMLCAGNSQRDSCGSPGVLCVKVGTWLQAGV 240
QY 241 VYMWDEGCQAQPNRPFGIYTRVTVYLLDWHIHVYVKPP 275
DB 241 VYMWDEGCQAQPNRPFGIYTRVTVYLLDWHIHVYVKPP 275

RESULT 12
AAW64238 standard; protein; 273 AA.

AC AAW64238;
XX 24-NOV-1998 (first entry)
DE Human mast cell tryptase I.
KW Mast cell tryptase I; human; MCP-7; mast cell protease 7; blood clot;
KW anticoagulant; myocardial infarction; reocclusion; thromboembolism;
KW cerebral embolism; thrombosis; therapy.
XX Homo sapiens.

OS
FH Key Location/Qualifiers
FT Peptide /label= Sig_peptide
FT Peptide /label= Pro_peptide
FT Peptide /label= Pro_peptide
FT Peptide /label= Mat_protein
XX MO9824886-A1.
XX 11-JUN-1998.
XX 25-NOV-1997; 97WO-US021620.
XX 04-DEC-1996; 96US-0032354P.
XX (BGHM) BRIGHAM & WOMENS HOSPITAL.

XX Stevens RL;
PI WPI; 1998-333308/29.
XX N-PSDB; AAW44329.
DR
DR
DR

New compositions containing tryptase-7, e.g. mouse mast cell protease-7 - are used to treat clot formation in e.g. myocardial infarction, reocclusion following angioplasty or pulmonary thrombo-embolism.

Dislosure; Page 65; 92pp; English.
This is the deduced amino acid sequence of human mast cell tryptase I (see also AAW44329). The invention provides compositions comprising an isolated tryptase-7 that may include chimeric proteins that contain (a) a human tryptase for all but the active site region and (b) the substrate-binding pocket of mouse tryptase-7 or its homologues (see AAW64233-39); a method for treating a blood clot by administering a nucleic acid molecule that codes for a tryptase-7, or an expression product, to decrease fibrinogen activity; a nucleic acid encoding a serine protease (SP) and a method of producing a mature SP by expressing the inactive zymogen in a host cell, and cleaving the enterokinase susceptibility domain. The tryptase-7 polypeptides can be used to treat disorders mediated by undesirable thrombus clot formation such as myocardial infarction and reocclusion following angioplasty of blood clots associated with pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal vein and peripheral arterial thrombosis. They are also useful for all surgical procedures that require decreased blood clots

Query Match 92.7%; Score 1402; DB 2; Length 273;
Best Local Similarity 92.3%; Pred. No. 1.9e-121;
Matches 252; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 3 SLLTLALPVLASRAVAAPVQALQOAGIVGGQEARPKMPOVSLRVRDRYMMHFCGGS 62
DB 1 MNLTLALPVLASRAVAAPVQALQOAGIVGGQEARPKMPOVSLRVHGPMWHFCGGS 60
QY 63 LIHPQWVLTAAHCIGRPDVKDLATLRVQLREONLHYODQLPVSRITVHPQFYIITQGADI 122
DB 61 LIHPQWVLTAAHCIGRPDVKDLALRVQLREONLHYODQLPVSRITVHPQFYIITQGADI 120
QY 123 ALLELEBEPVNISSRVHTVMLPPASERPPGMPGVDNDEPLRPPLKQVPI 182
DB 121 ALLELEBEPVNISSRVHTVMLPPASERPPGMPGVDNDEPLRPPLKQVPI 180
QY 183 MENHICDAKYNHLAGAVTGDDVRIIRDVMLCAGNSQRDSCGSPGVLCVKVGTWLQAGV 242
DB 181 MENHICDAKYNHLAGAVTGDDVRIIRDVMLCAGNSQRDSCGSPGVLCVKVGTWLQAGV 240
QY 243 SMDGEGQAQPNRPFGIYTRVTVYLLDWHIHVYVKPP 275
DB 241 SMDGEGQAQPNRPFGIYTRVTVYLLDWHIHVYVKPP 273

RESULT 13
AAW63174 standard; protein; 273 AA.

DE Human mast cell tryptase I polypeptide.
KW Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW tryptase-6 protease; inflammatory disorder; allergic rhinitis; urticaria;
KW antioedema; eczematous dermatitis; atopie dermatitis; anaphylaxis;
KW hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW inflammatory skin condition; human; mast cell tryptase I.
XX Homo sapiens.
OS

```

XX XX W09833812-A1.
XX PN
XX 06-AUG-1998.
XX PD
XX PF 30-JAN-1998; 98WC-US001865.
XX PR 05-FEB-1997; 97US-0037090P.
XX RR (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX PA
XX PI Stevens RL, Huang C;
XX PP WPI; 1998-4437390/37.
XX DR N-PSDB; AA42711.
XX XX
XX PT Trypsin-6 complex inhibitory peptides - used to treat mast cell-mediated
XX PT inflammatory disorders e.g. asthma.
XX PS Dislosure; Page 46-47; 69pp; English.
XX CC
XX CC This represents the human mast cell tryptase I protein which is a
XX CC homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention
XX CC provides sequences shown in AAW63160 to AAW63169 that are inhibitors of
XX CC mMCP-6. These peptides which are trypsin-6 complex inhibitors, can be
XX CC used for treating a mast cell-mediated inflammatory disorder. The
XX CC inhibitors can be used to treat inflammatory disorders including asthma,
XX CC allergic rhinitis, urticaria and anti-oedema, eczematous dermatitis
XX CC (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic
XX CC ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory
XX CC skin conditions
XX CC
XX SQ Sequence 273 AA;
  
```

Query Match 92.7%; Score 1402; DB 2; Length 273;

Best Local Similarity 92.3%; Pred. No. 1.9e-121; Mismatches 13; Gaps 0;

```

Matches 252; Conservative 8; Indels 0;
QY 3 SLLLLALPVLASRAVAAPAVQALQOAGIVGQGEAPRSKMPQVSLRVRDRYWMHFCGGS 62
Db 1 NULLLLPVLASRAVAAPAVQALQOAGIVGQGEAPRSKMPQVSLRVRDRYWMHFCGGS 60
QY 63 LHPQWVLTAAHCLGPDVYKDLATLRYQLRPHQLYYDDLLPVSRIIVHPQFYIIGTGADI 122
Db 61 LHPQWVLTAAHCLGPDVYKDLATLRYQLRPHQLYYDDLLPVSRIIVHPQFYIIGTGADI 120
QY 123 ALLELEPEPNVISRVTWMLPPRSSEPPRGMPCWMTVGMGDVNDDELPPRPLKQVYVPI 182
Db 121 ALLELEPEPNVISRVTWMLPPRSSEPPRGMPCWMTVGMGDVNDDELPPRPLKQVYVPI 180
QY 183 MENHICDAKXKHLGAVYTGDDVRIIRDDMLCAGNSQSDSCGSGSPVYCKVNGTMTLQAGVV 242
Db 181 MENHICDAKXKHLGAVYTGDDVRIIRDDMLCAGNSQSDSCGSGSPVYCKVNGTMTLQAGVV 240
QY 243 SWDEGCAQPNRPGIYRVRVYVYLDWIHHVYVKKP 275
Db 241 SWDEGCAQPNRPGIYRVRVYVYLDWIHHVYVKKP 273
  
```

```

RESULT 14
AAE14348
ID AAE14348 standard; protein; 691 AA.
XX
XX AAE14348;
AC
XX
DT 07-MAR-2002 (first entry)
DE Human protease PRPS-13 protein.
XX
XX
  
```

```

KW Human: protease: PRTS-13; tranquilliser; gene therapy; vaccine; allergy;
KW infection; dermatitis; arteriosclerosis; rheumatoid arthritis; hepatitis;
KW atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;
KW gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;
  
```

```

KW epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;
KW hypertension; neurological disorder; Parkinson's disease; drug screening;
KW cardiant; cell proliferative disorder; multiple sclerosis; osteoporosis;
KW diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;
KW autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;
KW developmental disorder; reproductive disorder; infertility; diarrhoea;
KW dementia; acidosis; cataract; gynaecomastia; epilepsy; jaundice.
XX
XX Homo sapiens.
OS
XX
FH Key
FH Peptide
FT 1..28
FT /label=Signal_peptide
FT 29..691
FT /note="Mature human protease PRTS-13 protein"
FT Active-site
FT 65..108
FT Domain
FT 67..80
FT /note="Type I fibronectin domain"
FT Active-site
FT 213..258
FT Domain
FT 225..238
FT /note="Type I fibronectin domain"
FT Domain
FT 371..389
FT /note="Transmembrane domain"
FT Active-site
FT 425..473
FT Domain
FT 433..450
FT /note="Kring1 domain protein"
FT Domain
FT 497..517
FT /note="Transmembrane domain"
FT 514..535
FT /note="Kring1 domain protein"
FT Active-site
FT 574..618
FT Domain
FT 594..635
FT /note="Kring1 domain protein"
FT Domain
FT 604..638
FT /note="Type I fibronectin domain"
XX
XX WO200183775-A2.
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014651.
XX
XX 04-MAY-2000; 2000US-0202082P.
XX 11-MAY-2000; 2000US-0203566P.
XX 17-MAY-2000; 2000US-0205803P.
XX 25-MAY-2000; 2000US-0207477P.
XX 01-JUN-2000; 2000US-0209402P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Delegeane AM, Lal P, Hafalia A, Patterson C, Walla NK, Kearney L;
XX Tribouley CM, Khan FA, Yao MG, Baughn MR, Azimzai Y, Elliott VS;
XX Nguyen DB, Gandhi AR, Yang J, Hernandez R, Policky JL, Lu DAM;
XX Reddy R, Yue H, Tang YT;
XX
XX WPI; 2002-034518/04.
XX N-PSDB; AAD23854.
XX
XX Novel human proteases and polynucleotides encoding the proteases, useful
XX for treating, diagnosing or preventing cell proliferative,
XX cardiovascular, autoimmune/inflammatory, neurological and developmental
XX disorders.
XX
XX Claim 1; Page 136-137; 151pp; English.
XX
XX The invention relates to human proteases (PRTS-14) and its corresponding
XX cDNA molecules. Human PRTS and its nucleic acid molecule are useful for
XX the diagnosis, treatment and prevention of disorders associated with
XX increased or decreased expression of PRTS. Examples of such disorders
XX include, cell proliferative disorders (arteriosclerosis, atherosclerosis,
XX hepatitis, psoriasis and cancer); autoimmune/inflammatory disorders
XX (AIDS, Addison's disease, allergy, anaemia, asthma, atopic dermatitis,
XX diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
  
```

```

  
```


trauma, Grave's disease, rheumatoid arthritis, ulcerative colitis, and viral, bacterial, fungal, parasitic, protozoal and helminthic infections) ; cardiovascular disorders (myocardial infarction, ischaemic heart disease and hypertension) ; neurological disorders (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease, stroke, mental disorders including mood, anxiety and seasonal affective disorder and prion diseases) ; gastrointestinal disorders (Crohn's disease, anorexia, nausea, diarrhoea and jaundice) ; epithelial disorders (contact dermatitis, eczema, acne vulgaris, alopecia, scabies, insect bites and urticaria) ; reproductive disorder (infertility) ; disruption of estrous and menstrual cycle and gynaecomastia) ; and developmental disorders (renal tubular acidosis, Cushing's syndrome, seizure disorders, congenital glaucoma and cataract) . PRTS DNA is also in useful as gene therapy. PRTS and its immunogenic fragments are useful for screening libraries of compounds in several drug screening assays. The present sequence is human protease PRTS-13 protein

Query Match 92.7% ; Score 1401 ; DB 5 ; Length 691 ; Best Local Similarity 92.0% ; Pred. No. 7, 6e-121 ; Mismatches 12 ; Gaps 1 ; Matches 253 ; Conservative 8 ; Indels 2 ; Gaps 1 ;

1 MSLLLALPVLASRAVAAPVQALQOAGIVGGQEARPSKWPQVSLRVRDRYMMHFCG 60
 11 MNLILLALPVLASRAVAAPVQALQOAGIVGGQEARPSKWPQVSLRVRDRYMMHFCG 68
 QY 61 GSLHPQWVLTAAHCLGPRVDKDLATLRVQLRBOHLYYODQLLPSRRIIVHPQFYIIQTGA 120
 DB 69 GSLHPQWVLTAAHCLGPRVDKDLALRVQLRBOHLYYODQLLPSRRIIVHPQFYIIQTGA 128
 QY 121 DIALLEBEPVNISSRVHTVMLPPASETFPPGMPCKVGTGMDVNDERLPPFPPLKQVVP 180
 DB 129 DIALLEBEPVNISSRVHTVMLPPASETFPPGMPCKVGTGMDVNDERLPPFPPLKQVVP 188
 QY 181 PIMENHICDAKYHLGAYTGDVRIIRDDMLCAGNSQRDSCKGDSCGSPILVCKVNGTWMLOAG 240
 DB 189 PIMENHICDAKYHLGAYTGDVRIIRDDMLCAGNTRRSDSCGSDSGPILVCKVNGTWMLOAG 248
 QY 241 VSWDEGCAQPNRPGIYTRVYTYLLDMVHHYVPKKP 275
 DB 249 VSWDEGCAQPNRPGIYTRVYTYLLDMVHHYVPKKP 283

RESULT 15
 AAM64240 standard; protein; 274 AA.
 ID AAM64240
 AC AAM64240;
 XX
 XX AAM64240;
 DT 24-NOV-1998 (first entry)
 XX
 DE Human mast cell tryptase II/beta.
 XX
 KM Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
 KM blood clot; anticoagulant; myocardial infarction; reocclusion;
 KM thromboembolism; cerebral embolism; thrombolysis; therapy.
 OS Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..18
 FT Peptide /label= Sig_peptide
 FT Peptide 19..28
 FT Peptide /label= Pro_peptide
 FT Peptide 29..274
 FT Peptide /label= Mat_protein
 XX
 PN MO9824886-A1.
 XX
 XX 11-JUN-1998.
 PD
 XX 25-NOV-1997; 97MO-US021620.
 PF

04-DEC-1996; 96US-0032354P.
 (BGMH) BRIGHAM & WOMENS HOSPITAL.
 Stevens RL;
 WPI: 1998-333308/29.
 N-PSDB: AAV44330.

This is the deduced amino acid sequence of human mast cell tryptase II/beta (see also AAV44330). The invention provides compositions comprising an isolated tryptase-7 that may include chimeric proteins that contain (a) a human tryptase for all but the active site region and (b) the substrate-binding pocket of mouse tryptase-7 or its homologues (see AAM64233-39); a method for treating a blood clot by administering a nucleic acid molecule that codes for a tryptase-7, or an expression product, to decrease fibrinogen activity; a nucleic acid encoding a serine protease (SP); and a method of producing a mature SP by expressing the inactive zymogen in a host cell, and cleaving the enterokinase susceptibility domain. The tryptase-7 polypeptides can be used to treat disorders mediated by undesirable thrombus clot formation such as myocardial infarction and reocclusion following angioplasty of blood clots associated with pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal vein and peripheral arterial thrombosis. They are also useful for all surgical procedures that require decreased blood clots

Query Match 92.6% ; Score 1400 ; DB 2 ; Length 274 ; Best Local Similarity 92.0% ; Pred. No. 3e-121 ; Mismatches 14 ; Indels 0 ; Gaps 0 ; Matches 252 ; Conservative 8 ; Mismatches 14 ; Indels 0 ; Gaps 0 ;

2 LSLLLALPVLASRAVAAPVQALQOAGIVGGQEARPSKWPQVSLRVRDRYMMHFCG 61
 1 INLLILLALPVLASRAVAAPVQALQOAGIVGGQEARPSKWPQVSLRVRDRYMMHFCG 60
 QY 62 SLHPQWVLTAAHCLGPRVDKDLATLRVQLRBOHLYYODQLLPSRRIIVHPQFYIIQTGA 121
 DB 61 SLHPQWVLTAAHCLGPRVDKDLALRVQLRBOHLYYODQLLPSRRIIVHPQFYIIQTGA 120
 QY 122 IALLEBEPVNISSRVHTVMLPPASETFPPGMPCKVGTGMDVNDERLPPFPPLKQVVP 181
 DB 121 IALLEBEPVNISSRVHTVMLPPASETFPPGMPCKVGTGMDVNDERLPPFPPLKQVVP 180
 QY 182 IMENHICDAKYHLGAYTGDVRIIRDDMLCAGNSQRDSCKGDSCGSPILVCKVNGTWMLOAG 241
 DB 181 IMENHICDAKYHLGAYTGDVRIIRDDMLCAGNTRRSDSCGSDSGPILVCKVNGTWMLOAG 240
 QY 242 VSWDEGCAQPNRPGIYTRVYTYLLDMVHHYVPKKP 275
 DB 241 VSWDEGCAQPNRPGIYTRVYTYLLDMVHHYVPKKP 274

Search completed: August 27, 2005, 23:26:46
 SDD time : 127 secs

This Page Blank (Just to)